Thu

```
updates/sec
                                                                                                                                         KITPQ 100
                                                                      Seconds
                                                                   ; Search time 38.6861 (without alignments) 287.115 Million cell
                                                                                                                                         .... DDFDLDVVKVSKQDS
                                                                                                                                                                                                                          747574
5.1.3
Compugen Ltd
                                                                                                                                                                                                                          of hits satisfying chosen parameters
                                                                                                                                                                                                   747574 segs, 111073796 residues
version
- 2002
                                                                                                                                                                                                                                                                                                           summaries
                                               model
                                                                     2002, 09:23:41
                                                                                                                                           1 MEKLFKEVKLEELENQKGSG...
                                                                                                                                                                 BLOSUM62
Gapop 10.0 , Gapext 0.5
                                                ΝS
 GenCore
(c) 1993
                                                                                                                                                                                                                                                                                 Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 su
                                                using
                                                                                                                                                                                                                                                  0
2000000000
                                                                                                             US-09-894-030-3
545
                                               search,
                                                                     October 24,
                                                                                                                                                                                                                                                 length:
length:
                                               protein
                                                                                                                                                                                                                                                  sed
                                                                                                                                                                 Scoring table:
                                                                                                                                score:
                                                                                                                                                                                                                           Total number
                                                                                                                                                                                                                                                 Minimum DB
Maximum DB
                                               protein
                                                                                                                                           Sequence:
                                                                                                                                                                                                     Searched:
                                                                                                                    Title:
Perfect
                                                                       Run on:
```

Database

icted by chance to have a sof the result being printed, score distribution. greater than or equal to the score of the store by derived by analysis of the Pred. No score gand is

# SUMMARIES

Description	o aci	Amino acid sequenc	Amino acid sequenc	TrpE/androgen rece	Human androgen rec	Human androgen rec	Human androgen rec	Human androgen rec	Androgen receptor.	Human androgen rec	Human gene 1-encod
ID	AAY03205	AAY03207	AAY03218	AAR12229	AAR12223	AAY33491	AAP93109	AAP90996	AAW14783	AAY78914	AAG71269
	20	20	20	12	12	20	10	10	18	21	22
Length DB	56	37	37	576	918	918	919	919	919	919	94
% Query Match	55.8	ω.	37.1	26.1	26.1	26.1	9	9	26.1	26.1	25.7
	04	211	202	42	CA	42	42	CA	42	142.5	140
Result No.		7	m	4	2	9	7	80	σ	10	11

Murine mV1 protein Murine BMP-12 homo	rtial m	n peptide #	#325 en	ein #	ain e	n bone ma	ide #308	ide #318	ide #299 enco	DNA sequence	n calpain s	an calpain	sophila mel	sophila mela	sophila me	dopsis tha	el fusion prot	ipoprotein	-7 C-terminal	n pr	protein S	n fibrillar	cancer a	de #2563	tide #2596 en	#2508	n brain ex	nan bone marro	eptide #2539 e	ptide #2633 en	ptide #2514 en	Apolipoprotein (a)
AAR78738 AAW26594	-	4	ABB32819	0	62	9		628	$\vdash$	'n	164	12	ABB58597	001	01	AAG51723	1856	363	œ	c	AAM79340	œ	391	991	509	ABB20509	59	28	10	859	AAM03832	AAW13632
16 18	22	22	22	22	22	22	22	22	22	19	18	22	22	22	22	21	18	18	16	22	22	15	21	22	22	22	22	22	22	22	22	18
240		0		0	0	0	101	101	101		9	9	118	δ	4	258	0	20	9	9	9	$^{\circ}$	$\boldsymbol{c}$	Ŋ	357	2	357	5	5	357	2	20
24.9	:	_:	_			_:			_:	m	œ.	Ψ.	w.	ω.	ω.	ω.	ω.	~	ď	ζ.	α.	ď	ď	ď	N.	ď	ď	ď	~	ď	ä	ď
135.5	35.	34.	34.	34.	34.	34.	34.	34.	34.	29.	12	$\sim$		12		12	$\sim$	$\sim$	$^{\circ}$	$^{\circ}$	12	24.	24.	24.	24.	24.	24.	24.	24.	•	24.	12
12	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45

### ALIGNMENTS

Æ

```
Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;
Gram-positive bacteria; pre-sublancin 168; sunA; sunT.
                                                                                                                                                                                                                                                                                                                         New antimicrobial peptide, sublancin 168, from Bacillus subtilis used for, e.g. treatment of infections caused by Gram negative bacteria and as food preservative
                                                                                    Amino acid sequence of sunA protein
                                                                                                                                                                                                                                                  (UYMA-) UNIV MARYLAND BALTIMORE
RESULT 1
AAY03205
ID AAY03205 standard; Protein;
                                                                                                                                                                                                       98WO-US14547
                                                                                                                                                                                                                            97US-0053035
                                                               (first entry)
                                                                                                                                                                                                                                                                                            WPI; 1999-131752/11.
N-PSDB; AAX28631.
                                                                                                                                        Bacillus subtilis
                                                                                                                                                              WO9903352-A1.
                                                                                                                                                                                                        17-JUL-1998;
                                                                                                                                                                                   28-JAN-1999.
                                                                                                                                                                                                                             18-JUL-1997;
                                                              03-AUG-1999
                                                                                                                                                                                                                                                                        Hansen JN;
                                         AAY03205;
```

d d

Óλ

```
Sublancin 168; antimicrobial; food preservative; pro-sublancin 168;
Gram-positive bacteria; Pro-sublancin 168; pre-sublancin 168.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             New antimicrobial peptide, sublancin 168, from Ba
used for, e.g. treatment of infections caused by
bacteria and as food preservative
     GLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (UYMA-) UNIV MARYLAND BALTIMORE
                                                                                                            standard; Protein;
                                                                                                                                                                      (first
                                                                                                                                                                                                 Amino acid sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1999-131752/11.
                                                                                                                                                                                                                                                                         subtilis
                                                                                                                                                                                                                                                                                                                   Disulfide-bond
Disulfide-bond
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Claim 13; Fig
                                                                                                                                                                                                                                                                                                                                                 Modified-site
                                                                                                                                                                                                                                                                                                                                                                              Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9903352-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              17-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   28-JAN-1999.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           18-JUL-1997;
                                                                                                                                                                      03-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                          Cross-links
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Hansen JN;
                                                                                                                                         AAY03218;
                                                                                                           AAY03218
                                                                                                                                                                                                                                                                         Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
      20
                                 ᆏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20
                                                                              RESULT 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Н
                                   Dp
                                                                                                                                                                                  d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                                                                                                                                                                                                            ö
                           This is the amino acid sequence of the sunA protein encoded by the prosublancin 168 nucleotide sequence used in the method of the invention. The peptide designated sublancin 168, is an antimicrobial useful for treating infections and preserving food against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years in aqueous solution of about
                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This is the amino acid sequence of prosublancin 168 used in the method of the invention. The peptide designated sublancin 168, is an antimicrobial useful for treating infections and preserving food against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years in aqueous solution of about neutral pH.
                                                                                                                       damage.
about
                                                                                                                                                                                                                                                                                                  56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Sublancin 168; antimicrobial; food preservative; pro-sublancin
Gram-positive bacteria; pre-sublancin 168.
                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                sublancin 168, from Bacillus subtinfections caused by Gram negative
                                                                                                      recursors of sublancin 168. Sub
and can be autoclaved without
2 years in aqueous solution of
                                                                                                                                                                                                              Length 56;
                                                                                                                                                                                                         re 304; DB 20;
red. No. 1.2e-24;
Mismatches 0;
                                                                                                                                                                                                           Score 304;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               prosublancin 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 le, sublancin lof info
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     English.
                                                                                                                                                                                                                         Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       and as food preservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MARYLAND BALTIMORE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      38.7%; Sc.
100.0%; P
                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                        37
71pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          antimicrobial;
                                                                                                                                                                                                           55.8%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Claim 1; Pages 52-53; 71pp;
                                                                                                                                                                                                                                                                                                                                                                        Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           98WO-US14547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97US-0053035
                                                                                                                                                                                                                                                                                                                                                                                                                                entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            antimicrobial peptide,
                                                                                                                                                                                                                       Similarity 100
6; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            treatment
                                                                                                                                                                                                                                                                                                                                                                                                                                                               of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WPI; 1999-131752/11.
                                                                                                                                                                                                                                                                                                                                                                                                                               (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                              sednence
                                                                                                                                                                                                                                                                                                                                                                       standard;
                                                                                                                                                                            56 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AA;
6; Fig 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sublancin 168;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            e.g.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (UYMA-) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WO9903352-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         17-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .8-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                                               03-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                             acid
                                                                                                                                                                                                                                       56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       used for,
                                                                                                                                                                                                         Query Match
Best Local S
                                                                                                                                                                             Sequence
                                                                                                                                                                                                                                                                                                                                                                   AAY03207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        bacteria
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                 neutral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Hansen
Claim
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Amino
                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                     207
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New
                                                                                                                                                                                                                                                                                                                                        RESULT
AAY032(
```

" (dda)

"dehydrogenated Thr

/note= 19..22 19..22 /note=

"thioether bridge"

97US-0053035

98WO-US14547

(Dha)"

Ser

"dehydrogenated

/note=

16

19

Location/Qualifiers7..36

7..36

of sublancin 168.

entry)

¥.

37

```
;
0
                                                                                                                                                                                                                                                                                     Gaps
                                                           is
food
                                                                                                                                  þe
                                 This is the amino acid sequence of sublancin 168 used in the method of the invention. The peptide designated sublancin 168, i an antimicrobial useful for treating infections and preserving against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years i aqueous solution of about neutral pH.
                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                          Score 202; DB 20;
Pred. No. 3.4e-14;
0; Mismatches 2;
                                                                                                                                                                                                                                                                                                                       56
                                                                                                                                                                                                                                                                                                                 AAR12229 standard; Protein; 576 AA
 71pp; English.
                                                                                                                                                                                                                                         37.1%;
ilarity 94.6%;
Conservative (
                                                                                                                                                                                                                                                          l Similarity
35; Conserv
7;
                                                                                                                                                                                                       37 AA;
                                                                                                                                                                                                                                                                                                                                                                                                            AAR12229
ID
```

ö

Gaps

Indels

Length

ore 211; DB 20; red. No. 3.9e-15; Mismatches 0;

Score 211; Pred. No.

Conservative

1 Similarity 37; Conserv

Query Match Best Local

Matches

Bacillus subtilis by Gram negative

```
рp
                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                 Ω
                                                                                                                                                                                                                                                                                        å
                                                                                                                                                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                    464
                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                      pATH
                                                                                                                                                                                                                                                                                                                                                               0000000
                                                                                                                                                                                                                                                             To express an androgen receptor fusion protein in E.coli, the pATI expression system was used. The trpE is insoluble so partially purified induced fusion protein is obtained by simply lysing the E.coli and precipitating the insoluble fusion protein. The fusion protein was used for immunisation to obtain monoclonal anti-AR antibodies.
                                                                                                                                                                                                                                                                                                                                          15;
                                                                                                                                                                                                                                                                                                                            576;
                                                                                                                                                                                                                                    and TR2 DNA binding proteins - DNA sequences detection and quantification methods
                                            DNA-binding protein; steroid hormone.
                                                                                                                                                                                                                                                                                                                                                              --666666666666
                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                          Indels
                                domain fusion protein
                                                                                                                                                                                                                                                                                                                       Score 142.5; DB 12;
Pred. No. 9.9e-07;
                                                                                                                                                                                                                                                                                                                             DB 12;
                                                                                                            AR
                                                                                                            region of
                                                                                               acid linker
                                                                                                                        5 amino acid linker
                                                                                  33kD trpE protein
                                                                                                                                                                                                                                                                                                                                                              416 GSGSPSAAASSSWHTLFTAEEGQLYGPCGGGG-
                                                                Location/Qualifiers
1..323
/label= 33kD trpE pro
324..329
                                                                                                            N-terminal
                                                                                                                                                                                                                                                                                                                                                                                                                             AA
                               TrpE/androgen receptor N-terminal
                                                                                   324..329
/label= 6 amino a
330..571
                                                                                                                                                                                                                                                      Example 13; Fig 9; 79pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                            918
                                                                                                                                                                                                                                                                                                                                          3,
                                                                                                                                                                                                                                                                                                                            26.1%;
ilarity 50.8%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                            Protein;
                                                                                                                                                              90WO-US06015
                                                                                                                                                                           89US-0438775
                   (first entry)
                                                                                                           /label= N
572..576
/label= 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Human androgen receptor
                                            androgen receptor; AR;
                                                                                                                                                                                         DEV CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                      (first
                                                                                                                                                                                                                 1991-178048/24
                                                                                                                                                                                                                                                                                                                                                                                                                            standard;
                                                                                                                                                                                                                                          antibodies for
                                                                                                                                                                                                                                                                                                                                    Similarity 32; Conser
                                                                                                                                                                                                                                                                                                                576 AA;
                                                                                                                                                                                                                                    Androgen receptor
                                                                                                                                                                                                     ΰ
                                                                                                                                                                                                                        N-PSDB; AAQ12007
                                                                                                                                                                                                     Chang
                                                                                                                                                                                       ARCH
                                                          sapiens
                                                                                                                                                                                                                                                                                                                                                                                            467
                                                                                                                                                                           17-NOV-1989;
                                                                                                                                                              19-OCT-1990;
                                                                                                                                    WO9107423-A
                                                                                                                                                                                                                                                                                                                                                                               76
                                                                                                                                                                                                                                                                                                                                                                                                                                                     20-AUG-1991
                   20-AUG-1991
                                                                                                                                                  30-MAY-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                         AAR12223;
                                                                                                                                                                                                                                                                                                                                                                                             999
                                                                                                                                                                                                                                                                                                                                                                                ggg
                                                                                                                                                                                                                                                                                                                                                                                                                            AAR12223
                                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                            Query Match
Best Local
                                                                                                                                                                                         (ARCH-)
                                                                                                                                                                                                     'n
                                                                           Protein
                                                                                                      Protein
                                                                                         Region
                                                                                                                  Region
                                                                                                                                                                                                     Liao
                                                                                                                                                                                                                                                                                                                                                       18
                                                                                                                                                                                                                                                                                                                                                                                            465
                                                         Ношо
                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                               74
                                                                                                                                                                                                                                           and
                                                                                                                                                                                                                                                                                                                                                                                                               23
                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
XEXEX
                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                            d
                                                                                                                                                                                                                                                                                                                                                       à
                                                                                                                                                                                                                                                                                                                                                                                ð
```

```
Proapoptotic; dependence domain; p75NTR; androgen receptor; DCC; huntingtin polypeptide; Machado-Joseph disease; SCA1; SCA2; SCA6; atrophin-1; cell death; apoptosis; Huntington's disease; head trauma; Alzheimer's disease; Kennedy's disease; spinocerebellar ataxia; stroke; dentatorubropallidoluysian atrophy; cell proliferation; cell survival; neoplastic; malignant; autoimmune; fibrotic.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GSGLGKAQCAALW--LQCASGGTI--GCGGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               This sequence was deduced from a cDNA clone isolated by screening commercially available human testis and prostate lambda gtll cDNA libraries. The sequence is very similar to that of rat AR and in the DNA-binding domain it is identical to that of rAR DNA-binding domain. Homology comparisons with other known steroid receptors indicate that hAR is more closely related to glucocorticoid, mineralo-corticoid and progesterone receptors than to v-erb-A or treceptors for oestrogen, vitamin D and thyroid hormones.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1111111111111111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and TR2 DNA binding proteins - DNA sequences detection and quantification methods
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 918;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142.5; DB 12;
Pred. No. 1.6e-06;
3; Mismatches 13;
                                                                                                                   domain
 steroid hormone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GSGSPSAAASSSWHTLFTAEEGOLYGPCGGGG---
                                                                                                                DNA-binding day Cysteine-rich
                                                                   Location/Qualifiers
556..626
/label= DNA-binding (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Human androgen receptor protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26.1%;
illarity 50.8%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               standard; Protein;
                                                                                                                                                                                                                                                       90WO-US06015
                                                                                                                                                                                                                                                                                          89US-0438775
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (first entry)
hAR; DNA-binding protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79pp;
                                                                                                                                                                                                                                                                                                                                  (ARCH-) ARCH DEV CORP
                                                                                                                                                                                                                                                                                                                                                                                                            1991-178048/24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              l Similarity
32; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Androgen receptor a and antibodies for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              918 AA;
                                                                                                                                                                                                                                                                                                                                                                         ပ်
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                               N-PSDB; AAQ12001
                                                                                                                                                                                                                                                                                                                                                                     Chang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Fig
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 W09945944-A1
                                         sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                468
                                                                                                                                                                                                                                                    19-OCT-1990;
                                                                                                                                                                                                                                                                                          17-NOV-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9-JAN-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9/
                                                                                                                                                                         WO9107423-A
                                                                                                                                                                                                                 30-MAY-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAY33491;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ggg
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  - B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Androgen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAY33491
                                                                                                                                                                                                                                                                                                                                                                         Liao S,
                                                                                                  Domain
                                       Ношо
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18
                                                                                                                                                                                                                                                                                                                                                                                                         WPI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             466
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   417
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 6
```

<u>ښ</u>

Lubahn DB;

Joseph DR,

```
Wilson EM,
                               1989-324206/44
                                             N-PSDB; AAN91772
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Chang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-MAR-1989;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     30-MAR-1988;
                                                                                                                     Disclosure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28-FEB-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             WO8909223-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          05-OCT-1989
FS,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAP90996;
                                                                                                                                                                                                                                                                       Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAP90996
                                                                                                                                                                                                                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Liao S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Region
                                                                                                                                                                                                                                                                                                                                                                                                                                                  443
                                                                                                                                                                                                                                                                                                                                                                                                                      55
                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                        386
                                WPI;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP90996
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                        Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             XX XX
                              g
                                                                                                                                                                                                                                                                                                                                                             ó
                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FT
FT
XX
                                                                                                                                                                                                                                                                   This invention describes novel pure proapoptotic dependence peptides
which comprise a sequence of an active dependence domain selected from
dependence polypeptides consisting of p75NTR, androgen receptor, DCC,
huntingtin polypeptide, Machado-Joseph disease gene product, SCA1, SCA2,
SCA6 and atrophin-1 polypeptide. The proapoptotic peptides are capable
of inducing cell death and can be used to develop products to mediate or
inhibit apoptosis. The methods can be used for reducing the severity of
a proapoptotic dependence domain mediated pathological conditions e.g.
Huntington's disease, Alzheimer's disease, Kennedy's disease,
Spinocerebellar ataxias, dentatorubropallidoluysian atrophy,
Machado-Joseph disease, stroke or head trauma. They can also be used for
reducing the severity of a pathological condition mediated by upregulated
cell proliferation or cell survival e.g. neoplastic, malignant,
autoimmune or fibrotic conditions. This sequence represents a human
capable described in the method of the invention.
                                                                                                                                                                                                                                                                                              ., sCA1, SCA2, class are capable ducts to mediate or ng the severity of a conditions e reserved.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4;
                                                                                                                                                                                                                                                                                                                                                                                                                     be used for upregulated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                               for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ts
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5;
                                                                                                                                                                                             New proapoptotic dependence peptides, used to develop productreating, e.g. Alzheimer's disease -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -GTIGCGGG---AVACQNYRQF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Human androgen receptor; ployclonal antibody; cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 142.5; DB 20;
Pred. No. 1.6e-06;
3; Mismatches 22;
                                                                                                                                                                                                                                        Disclosure; Page 90-93; 199pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GUYNC-) UNIVERSITY OF NORTH CAROLINA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --- CRGGGGGGGGGGGGGGGGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAP93109 standard; protein; 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ELENQKGSGLGKAQCAALWLQCASG
                              99WO-US05250.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local Similarity 42.5%;
Matches 37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        88US-0182646
                                                          98US-0041886
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            89WO-US01548
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                      ŝ
                                                                                                                      Rabizadeh
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Human androgen receptor
                                                                                        BURNHAM INST
                                                                                                                                                WPI; 1999-561617/47
N-PSDB; AAZ23424.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            918 AA;
                                                                                                                     Bredesen DE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    sapiens
                            11-MAR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3-APR-1989;
                                                          12-MAR-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WO8909791-A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14-APR-1988;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19-MAR-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              19-0CT-1989
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              AAP93109;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                       (BURN-)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   AAP93109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
```

a

à

qq

ŏ

ü

PN XX

```
Androgen receptor protein (AR) is used to produce mono- or poly-clonal antibodies. These are used for the detection and quantification of AR in the presence of endogenous androgen, as androgen will not interfere with binding. They may be used in assays to determine and quantify cellular distribution of AR in tumour tissue, and are esp. useful for evaluating distribution of AR in tumour tissue, and are esp. useful for evaluating
DNA encoding androgen receptor protein - useful for transforming eukaryotic hosts for protein expression and subsequent antibody prodn.
                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                             54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            encoding new androgen receptor and TR2 polypeptide(s) - ab DNA, and derived antibodies, useful for receptor assay and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       KLENPLDYGSAWAAAAA - - - QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                   25;
                                                                                                                                                                                                                                                                                                                                      Length 919;
                                                                                                                                                                                                                                                                                                                                                                                                                          --GTIGCGGG---AVACQNYRQF--
                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                      DB 10;
                                                                                                                                                                                                                                                                                                                                   Score 142.5; DB 10;
Pred. No. 1.6e-06;
; Mismatches 22;
                                                                                                                                                                                                                          prostate cancers to determine responsiveness therapy.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Location/Qualifiers
1..919
./*tag= a
/product=98 kD polypeptide
185..919
/*tag= b
/product=79 kD polypeptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Androgen receptor; TR2 polypeptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A
                                                                   41pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -CRGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               919
                                                                                                                                                                                                                                                                                                                                                                                 3;
                                                                                                                                                                                                                                                                                                                                                                                                                              12 ELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (ARCH-) ARCH DEVELOPMENT CORP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human androgen receptor DNA
                                                                                                                                                                                                                                                                                                                                   th 26.1%;
Similarity 42.5%;
37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               standard; protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           89WO-US01238
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      88US-0176107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (first entry)
                                                                 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          WPI; 1989-309501/42
N-PSDB; AAN91577.
                                                                                                                                                                                                                                                                                             919 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ပ်
                                                                 Fig.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            New DNA
to bind
```

54

-AVACQNYRQF-

GTIGCGGG-

.rag

g

5

ð

g

```
ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                  WO200001813-A2
                                                                                                                                                                                                                                                               sapiens.
                                                                                                                                                                                                                                                                                                                                                                         (UYBR-) UNIV
                                                                                                                                                                                                                                                                                                                                                   30-JUN-1998;
                                                                                                                                                                                                                                                                                                                              30-JUN-1999;
                                                                                                                                                                                                                                                                                                        13-JAN-2000
                                                                                                                                                                       23-MAY-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                               Sadar MD,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence
                                                                                                                                                  AAY78914;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              tumours
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                      WPI;
                                                                                                                                                                                                                                                               Homo
                            386
                                                                       443
                                                                                                        10
      12
                                                  55
                                                                                                        RESULT 10
AAY78914
                                                                                                                                                                                                                                                   XX DE XX KW KW KW
                                                                                                                                                   XX
                                                                       qq
                            g
                                                  Q
         ŏ
                                                                                                                                         4;
                                                                                                                                                                                    439
                                                                                                                                          Gaps
                                                                                                                                                              54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            pression,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Oligonucleotide(s) antisense to human androgen receptor and acidic FGF genes - used to inhibit gene expression, for the treatment of
                                                               product
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   acting
                                                                                                                                                                                    KLENPLDYGSAWAAAAA - - - QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                       therapy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Human androgen receptor (AAW14783) binds testosterone and, acting at the transcriptional level, regulates the growth of normal prostatic cells. Antisense oligonucleotides (see also AAT63200, AAT63404-05) based on an androgen receptor cDNA clone (see also AAT63407) can be used to prevent androgen receptor gene expressic thereby inhibiting the growth or survival of prostatic cells for the treatment of benign prostatic hyperplasia and prostate cancel
                                         and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               919;
                                                                                                                                         25;
                                                                                                                    919;
                                            assays
                                                              ΚD
                                                                                                                                                                -GTIGCGGG---AVACQNYRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                            Androgen receptor; acidic fibroblast growth factor; aFGF; antisense; benign prostatic hyperplasia; prostate cancer;
                                                                                                                                          Indels
                                                              79
                                          The protein is used to raise antibodies for receptor
                                                                the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 18;
                                                                                                                   Score 142.5; DB 10;
Pred. No. 1.6e-06;
3; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 142.5; DB 18
Pred. No. 1.6e-06;
3; Mismatches 22;
                                                               cogon;
                                                              the first Met
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (WORC-) WORCESTER FOUND BIOMEDICAL RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                51pp; English.
                                                                                                                                                                                                                                 466
                                                                                                                                                                                                                                                                                       ¥
                                                                                                                                                                                                                     QLYGPCGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                            -crecececececececece
                                                                                                                                                                                                                                                                                       919
                                                                                                                                                                ELENOKGSGLGKAQCAALWLQCASG-
                                                                                                                                          3;
                  60pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         benign prostatic hyperplasia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26.1%;
42.5%;
                                                                                                                    26.1%;
ilarity 42.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  95US-0004018.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            96WO-US15081
                                                                at
                                                                                                                                                                                                                                                                                       Protein;
                                                                                                                                                                                                                                                                                                                                   entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 42.3
37; Conservative
                                                    affinity purification.
The 98 kD product starts
starts from the second.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Disclosure; Page 22-28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    WPI; 1997-202879/18
                                                                                                                                                                                                                                                                                                                                 (first
                                                                                                                             1 Similarity 37; Conserv
                                                                                                                                                                                                                                                                                       standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            919 AA;
                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                        receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAT63407
                      8; Fig 3;
                                                                                                919
                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                   WO9711170-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            20-SEP-1996;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   20-SEP-1995;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Zamecnik PA;
purification
                                                                                                                                                                                                                                                                                                                                 22-JUN-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                        27-MAR-1997
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local S
                                                                                                                                                                                                                                                                                                             AAW14783;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence
                                                                                                                                                                                                                                                                                                                                                        Androgen
                                                                                                                                                                                                                                                                                      AAW14783
                                                                                                 Sequence
                                                                                                                      Query Match
Best Local
                      Claim
                                                                                                                                  Best Loc
Matches
                                                                                                                                                                12
                                                                                                                                                                                      383
                                                                                                                                                                                                                                  440
                                                                                                                                                                                                                                                                 RESULT 9
```

PD XX

```
This sequence represents the human androgen receptor (AR) amino acid sequence. The invention relates to a fragment of the AR corresponding to amino acids 234-391 (see AAY78913). The fragment mediates amino acids 234-391 (see AAY78913). The fragment mediates androgen-independent activation of the AR. The androgen receptor acts as androgen-responsive genes. Interaction of the AR with the protein kinase independent region. The AR fragment and peptides derived from it can be used as agents for inhibiting androgen independent activation of the androgen receptor, as activation domains, and as a tool for screening androgen receptor, as activation domains, and as a tool for screening for compounds which affect androgen-independent activation of the AR. The peptides, when used in combination with androgen deprivation, effectively limit androgen mediated disease progression. These diseases include cancer, benign prostatic hyperplasia, hirusutism, androgenic cancer. The peptides and nucleic acids encoding them, are especially used the propriet in an inventive in relational disease, especially prostate the primer in relational diseases, especially prostate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Androgen receptor; AR; androgen-independent activation; inhibitor; cancer; benign prostatic hyperplasia; hirusutism; androgenic alopecia; acne; breast cancer; Kennedy disease; prostate cancer.
                442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          diseases
KLENPLDYGSAWAAAAA --- QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -GTIGCGGG---AVACQNYRQF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Novel non-androgen ligand binding peptides for inhibiting androgen-independent activation of androgen receptor, used screening compounds and for treatment of androgen-mediated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mawji
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Human androgen receptor (AR) amino acid sequence.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142.5; DB 21
Pred. No. 1.6e-06;
; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Snoek
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               deprived of androgen.
                                                                                                                                   469
                                                                               26
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Disclosure; Page 7; 32pp; English.
                                                                                                                                                                                                                                                                                      Æ
                                                                                                            CRGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3;
                                                                                                                                                                                                                                                                                   AAY78914 standard; protein; 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gout
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              BRITISH COLUMBIA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          98US-0091871.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26.1%;
42.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              99WO-CA00604.
                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bruchovsky N,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              such as prostate cancer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                2000-182113/16.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 in patients
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 Similarity
37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           919 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
```

4;

4;

Gaps

25;

Indels

3;

Matches

```
8888888888888888
                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                              DE XX XX XX OS XX XX EH FT FT FT FT FT FT FT FT FT
                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FT
FT
FT
FT
FT
FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      presence of each of the y expressed,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 for
                                                                                                                                                                                                                                                  Human; secreted protein; proliferative disorder; cancer; chromosome 15; foetal abnormality; developmental abnormality; haematopoietic disorder; immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis; inflammation; allergy; neurological disorder; Alzheimer's disease; Parkinson's disease; cognitive disorder; schizophrenia; asthma; skin disorder; psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder; angiogenic disorder; kidney disorder; gastrointestinal disorder; pregnancy-related disorder; tumour; endocrine disorder; infection; wound healing; vulnerary; chemotaxis; food additive; binding partner identification.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               encode
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       inflammation,
                    442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         in or gene
Ining the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         endocrine
to aid wound
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 secreted
they enco
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      eventing,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 useful
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            lerosis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New nucleic acid molecule encoding a human secreted protein, usefu. preventing, treating or ameliorating medical conditions such as rheumatoid arthritis, Alzheimer's disease and microbial infections
                   FTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         developmental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     AAH31349-AAH31428 represent cDNAs corresponding to 26 human secret protein genes, and AAG71243-AAG71319 represent the proteins they and AAG71243-AAG71319 represent the proteins they and their corresponding secreted proteins are useful for preventities treating or ameliorating medical conditions, e.g., by protein or therapy. Pathological conditions can be diagnosed by determining the presentations in the new protein in a sample or by determining the present mutations in the new genes. Specific uses are described for each and include developing products for the diagnosis or treatment of proliferative disorders, cancer, tumours, foetal and developmental
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   proliferative disorders, cancer, tumours, foetal and developmenta abnormalities, haematopoietic disorders, diseases of the immune s. AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammat allergies, neurological disorders (e.g., Alzheimer's disease, Parkinsons's disease), cognitive disorders, schizophrenia, asthma skin disorders (e.g., psoriasis), sepsis, diabetes, atheroscleros cardiovascular disorders, angiogenic disorders, kidney disorders, gastrointestinal disorders, pregnancy-related disorders, endocring disorders, and infections. The proteins can also be used to aid we
                                                                                                                                                                                                                            SEQ ID NO:118.
Moore PA;
                                                                                                                                                                                                                         Human gene 1-encoded secreted protein HTEGI42,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Young PE,
                                                                        469
                                           9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  581pp; English.
                                                          QLYGPCGGGGGGGGGGGGGGGGG
                                           -CRGGGGGGGGGGGGGGGG
                                                                                                                                      AAG71269 standard; Protein; 94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29-OCT-1999; 99US-0162211.
30-JUN-2000; 2000US-0215138.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2000WO-US29360.
                                                                                                                                                                                               entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Komatsoulis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (HUMA-) HUMAN GENOME
                                                                                                                                                                                              (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2001-291051/30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Claim 11; Page 496;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              N-PSDB; AAH31376.
                                                                                                                                                                                                                                                                                                                                                                                                                                                  WO200132674-A1
                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25-OCT-2000;
                                                                                                                                                                                              30-JUL-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             10-MAY-2001
                                                                                                                                                                   AAG71269;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ruben
             386
                                                                    443
                                        55
                                                                                                            RESULT 11
AAG71269
                                                                   q
                                                                                                                                                                                                                      δŽ
                                                                                                                                                                              XX
```

```
Sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues, to identify their cognate ligands or binding partners, and in chemotaxis, and can be used as a food additive or preservative to modify storage properties.

Antibodies specific for a protein of the invention can be used in alleviating symptoms associated with the disorders mentioned above, and in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked immunosorbent assay (ELISA). The present sequence represents a human secreted protein of the invention.
   due
                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                          14;
                                                                                                                                                                              94;
   skin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Wolfman NM;
                                                                                                                                                                                                         Indels
                                                                                                                                                                              Length
                                                                                                                                                                                                                                    79
   to prevent
                                                                                                                                                                                                                                 -666666666666666666666
                                                                                                                                                                                                                                                                                                                                                                                                                                          tendon; ligament
                                                                                                                                                                                                          .,
4
                                                                                                                                                                              DB 22;
2.7e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Thomsen GH,
                                                                                                                                                                             Score 140; DB
Pred. No. 2.7e
1; Mismatches
proliferation,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                240 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Rosen VA,
                                                                                                                                                                                                                                                                                                                                                                                                                                           mV1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Τ,Α
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         V, A, E, G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          \alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ß
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        V, A, E, G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        L, P, Q,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             V,A,D,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q, K, E
and epithelial cell
                                                                                                                                                                             25.7%;
                                                                                                                                                                                                                                                                                                                               standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           94US-0333576.
93US-0164103.
94US-0217780.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    S,P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Ω,Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                           protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  94WO-US14030
                                                                                                                                                                 Query Match
Best Local Similarity 59.6
Watches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /label-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         'label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HARVARD COLLEGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INST
                                                                                                                                                                                                                                                                                                                                                                                    (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Melton
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1995-224320/29
)B; AAQ96223.
                                                                                                                                                                                                                                                                                                                                                                                                                                          Bone morphogenetic
                                                                                                                                                                                                                                                                                                                                                                                                              Murine mV1 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     56
                                                                                                                                                                                                                                                 GENETICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         02-NOV-1994;
07-DEC-1993;
25-MAR-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                06-DEC-1994;
                                                                                                                                                                                                                                                                                                                                                                                   23-NOV-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           WO9516035-A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      15-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AJ,
JM;
                                                                                                                                                                                                                                                                                                                                                         AAR78738;
                                                                                                                                                                                                                                                                                                                              AAR78738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                WPI; 199
N-PSDB;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Celeste
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Wozney
                                                                                                                                                                                                                                 33
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (HARD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 (GEMY
                                                                                                                                                                                                                                                                                                    RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mus
                                                                                                                                                                                                                                                                                                               AAR78738
```

.rag

-09-894-030-3

```
07-DEC-1993;
                                                  22-DEC-1994;
07-DEC-1993;
25-MAR-1994;
02-NOV-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18-DEC-2001
                                                                                                                      AJ,
JM;
US5658882-A
                 19-AUG-1997
                                                                                                                                                                                                                                                                                                                                 36;
                                                                                                                                                                                                                                                                                                                 Query Match
Sest Local S
                                                                                                                                                                                                                                                                                                 Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                AAE10984
                                                                                                                       Celeste
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Mus sp
                                                                                                                               Wozney
                                                                                                     (HARD
                                                                                              (GEMY
                                                                                                                                                                                                                                                                                                                                                                                    51
                                                                                                                                                                                                                                                                                                                                                   7
                                                                                                                                                                                                                                                                                                                                                                    53
                                                                                                                                                                                                                                                                                                                                                                                                    107
                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                       AAE10984
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   KW
KW
KW
KW
KW
KW
KW
COS
COS
FH
FT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                 XEXTX
                                                                   d
                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                    Q
 PN
XX
PD
                         XX
PF
PR
                                                            PR
                                                                                                                                                                          4;
                                                                                                                                                                                                            106
                                                                                                                                                                           Gaps
                                                                                                                                                                                           50
                                           Oligos #6 and #7 (AAQ96218 & AAQ96219) are used as primers for the amplification of a 275 bp DNA probe, the internal 269 bp of which corresp. to nts #607 to #865 of AAQ96207, from the BMP-12 encoding plasmid subclone PCR1-1#2. This probe was radioactively labelled and used to screen a murine genomic library. DNA sequence analysis of one of positively hybridising recombinants named MVR3 indicates that it encodes a portion of the mouse gene corresp. to the PCR product mV1 (murine homolog of the human BMP-12 sequence AAQ96207. The partial DNA sequence of this subclone and corresp. AA translation are given in AAQ96223 & AAR78738.
  in
                                                                                                                                                                                           --ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN
                                                                                                                                                                                                                                                                                                                                                           BMP-12; bone morphogenetic protein; mouse; tendon; ligament; wound healing; tissue repair; carpal tunnel syndrome; tendonitis,
  nseq
                                                                                                                                                                          23;
                                                                                                                                                         240;
                                                                                                                                                                                                      roteins -12 and -13 and corresp. DNA -tendon/ligament-like tissue formation
                                                                                                                                                          Length
                                                                                                                                                                           Indels
                                                                                                                                                         24.9%; Score 135.5; DB 16;
40.4%; Pred. No. 2.1e-06;
ive 5; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CNG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ala
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Asp, (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                               "encoded by NCG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . Cys, Trp
"encoded by TGN"
                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= Val, Ala, Glu, /note= "encoded by GNG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "encoded by AGN"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gln,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Glu,
                                                                                                                                                                                                                            Thr.
                                                                                                                                                                                                                                                                                                                                                                                                                Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 /label= Lys, Gln
/note= "encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         val, Ala, C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     "encoded by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       note= "encoded by
                                                                                                                                                                                                                                                                                          240 AA
                            Example; Page 68-70; 84pp; English
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pro,
                                                                                                                                                                                                                                                                                                                                             Murine BMP-12 homologue fragment.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pro,
                                                                                                                                                                          5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ser,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Leu,
                                                                                                                                                                                                                                                                                          standard; Protein;
   morphogenetic proteins in. for inducing tendon,
                                                                                                                                                                                                                                                                                                                            (first entry)
                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                          /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                'label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      /note-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    'note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             note=
                                                                                                                                                                                             2 EKLFKEVKLE-----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Misc-difference 103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              26
                                                                                                                                                                   Local Similarity
nes 36; Conserv
                                                                                                                                          240 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                        Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                musculus
                                                                                                                                                                                                                                                                                                                            21-JAN-1998
                                                                                                                                                                                                                                                                                                            AAW26594;
                                                                                                                                                                                                                                                                                          AAW26594
                                                                                                                                           Sequence
                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                               therapy
                                                                                                                                                                                                                                              107
                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                         RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                Mus
                                                                                                                                                                                                                                                                                                                                                                                                                 Key
                                                                                                                                                                                                                                                                                  AAW26594
                                                                                                                                                                     Best
                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                g
   ò
```

```
Mouse; bone morphogenic protein; BMP-12; vulnerary; antiinflammatory; analgesic; gene therapy; transforming growth factor-beta; TGF-beta; tissue formation; wound healing; tissue repair; ligament defect; carpal tunnel syndrome; tendonitis; mV1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          50
                                                                                                                                                                                                                                                                                                                                                                                                      This polypeptide comprises a fragment of a murine homologue of human bone morphogenetic protein 12 (BMP-12) (see also AAW26589). Its amino acid sequence was deduced from DNA subclone mV1 (see AAT90396), isolated from murine genomic DNA using primers (see AAT90393-94) based on human BMP-12 sequences. Human BMP-12, BMP-13 (see AAW26591) and MP52 (see AAW26590) polypeptides are used in a claimed method for inducing tendon and ligament formation.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN
                                                                                                                                                                                                                                                                                               BMP-13 or M
tendonitis,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wolfman NM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                           Inducing tendon and ligament formation using BMP-12, useful for tissue healing and repair, treatment of improving fixation of tendons to bone etc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 135.5; DB 18;
Pred. No. 2.1e-06;
                                                                                                                                                                                      Thomsen GH,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Encoded by NCG"
                                                                                                                                                                                                                                                                                                                                                                          Example 1; Column 65-68; 43pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ρλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -GGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      240 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       YRQFCRGGGGGGGGGGGGGGGGGGMSK
                                                                                                                                                                                       Rosen VA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          "Encoded
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5;
                                    94US-0362670.
93US-0164103.
94US-0217780.
94US-0333576.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAE10984 standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Murine partial mV1 protein.
93US-0164103
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                entry)
                                                                                                                                S INST INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               'note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        /note=
                                                                                                                                                                                       Melton DA,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (first
                                                                                                                                                                                                                                          WPI; 1997-424270/39.
N-PSDB; AAT90396.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EKLFKEVKLE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 240 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
                                                                                                                                ) GENETICS
) HARVARD (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Key
Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Misc-difference
```

-09

us

-GGGGGGGGGGGGGGGAGR 129

107

acid probe

single exon nucleic

cell

breast

ρλ

encoded

#300

peptide

Human

(first entry)

01-FEB-2002

ABB27649

AA

101

Peptide;

ABB27649 standard;

RESULT 15 ABB27649

gene expression; breast;

exon probe;

single

microarray;

Human;

FT FT FT

PN

disease; cancer.

sapiens.

Ношо

WO200157271-A2

2001WO-US00662

30-JAN-2001;

09-AUG-2001.

04-FEB-2000;

```
The invention relates to a chimeric DNA comprising a LNA SCALLE.

C encoding a propeptide from a member of the transforming growth factor

C TGF)-beta superfamily of proteins. This DNA is linked to a DNA sequence

C of Bone Morphogenetic Protein (BMP)-12, BMP-13 or MP52 protein. The DNA

C sequences are useful for producing proteins which induce tendon/ligament

C sequences encoding BMP-12 related proteins with similar activity. The

S sequences encoding BMP-12 related proteins with similar activity. The

proteins are useful for the induction of tendon/ligament-like tissue

formation, wound healing, ligament and other tissue repair, augmenting

the activity of bone morphogenetic proteins, and for treating tendonitis,

carpal tunnel syndrome and other tendon and ligament defects. The

present sequence is murine partial mV1 protein which is homologous to

human BMP-12 or VL-1 sequences of the invention.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      encoding (BMP)-12,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     nnel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 tendonitis, carpal tunn
defects, comprises DNA
morphogenetic proteins
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Thomsen GH;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wolfman NM,
                                                                                                  CNG"
                                                                                                                                                      TGN"
                                                                                                                                                                                                             GNC"
                                                                                                                                                                                                                                                                   GNG"
                                                                                                                                                                                                                                                                                                                          by NAG"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          New chimeric DNAs, useful for treating syndrome and other tendon and ligament propeptide linked to DNA encoding bone BMP-13 or MP52 -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         English
                                         рy
                                                                                                рy
                                                                                                                                                      \mathbf{p}
                                                                                                                                                                                                             ρλ
                                                                                                                                                                                                                                                                  by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Rosen VA,
                                        "Encoded
                       Unknown
                                                                                                                                                   "Encoded
                                                                                                                                                                                                                                                                  Encoded
                                                                                                                                                                                                                                                                                                   /label= Unknown
/note= "Encoded
                                                                            Unknown
                                                                                            "Encoded
                                                                                                                                  Unknown
                                                                                                                                                                                                             Encoded
                                                                                                                                                                                           Unknown
                                                                                                                                                                                                                                                 Unknown
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     42pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   94US-0362670.
93US-0164103.
94US-0217780.
94US-0333576.
                                                                                                                                                                                                                                                                                                                                                                                                                                  97US-0808324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GENETICS INST INC
HARVARD COLLEGE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Example 1; Column 67-70;
                        'label=
                                                                                                                                 label=
                                                                                                                                                                                         label=
                                                                                                                                                                                                                                               /label=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wozney JM,
                                         note=
                                                                                              note=
                                                                                                                                                                                                                                                                 'note=
                                                                                                                                                                                                             note=
                                                                                                                                                                                                                                                                                                                          /note=
                                                                            label
                                                                                                                                                      note=
                                                                                                                                                                                                                                                                              104
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       WPI; 2001-588978/66
N-PSDB; AAD18335.
                                                                                                                                                                                                                             10
Misc-difference
                                                       Misc-difference
                                                                                                             Misc-difference
                                                                                                                                                                      Misc-difference
                                                                                                                                                                                                                          Misc-difference
                                                                                                                                                                                                                                                                              Misc-difference
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22-DEC-1994;
07-DEC-1993;
25-MAR-1994;
02-NOV-1994;
                                                                                                                                                                                                                                                                                                                                                                                                                               28-FEB-1997;
                                                                                                                                                                                                                                                                                                                                                         US6284872-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AJ,
                                                                                                                                                                                                                                                                                                                                                                                             04-SEP-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Celeste
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Melton
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (GEMY
(HARD
```

```
The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human breast and BT 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for determining predisposition and/or prognosing breast disease. Gene expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The probes of the invention acid as single exon nucleic acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GGGGGGGG 47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KEVKLEELENQKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                       exon nucleic acid probes, sample derived from human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 101;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Note: The sequence data for this patent did not form part of printed specification, but was obtained in electronic format from WIPO at ftp.wipo.int/pub/published_pct_sequences.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              English
                                                                                                                                                                                                                                                                                                                                                                                                                                                 nucleic acid probes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence listing;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               15;
15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 134.5; DB : Pred. No. 1.1e-06
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                - - 0000000000--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                     set of single expression in sof single exon r
                                                                                                                                                                                                                                                                                                                                                                  Rank DR;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              +
                                                                                                                                                                                                                                                                                                                                                                                                                    of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           327pp
                                                                                                                                                                                                                                                                                                                                      (MOLE-) MOLECULAR DYNAMICS INC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                                                                     spatially-addressable set
                                                                                                                                                                                                                                                                                                                                                                  3
                                                                                                                                                                                                                        2000US-0180312.
2000US-0207456.
2000US-0608408.
2000US-0632366.
2000US-0234687.
2000US-0234587.
                                                                                                                                                                                                                                                                                                           2000GB-0024263
                                                                                                                                                                                                                                                                                                                                                                                                                                    useful for measuring gene ex
breast, comprises number of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               24.78;
44.68;
                                                                                                                                                                                                                                                                                                                                                                Chen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SEQ ID NO 10617;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KRKKKEEEEGREEEG-GR--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 the invention.
                                                                                                                                                                                                                                                                                                                                                               Hanzel DK,
                                                                                                                                                                                                                                                                                                                                                                                          WPI; 2001-496933/54.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity 33; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   101 AA;
                                                                                                                                                                                                                                    26-MAY-2000;
30-JUN-2000;
03-AUG-2000;
21-SEP-2000;
27-SEP-2000;
                                                                                                                                                                                                                                                                                                        04-OCT-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 27;
                                                                                                                                                                                                                                                                                                                                                                SG,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              probe of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local (
                                                                                                                                                                                                                                                                                                                                                                                                                                useful
                                                                                                                                                                                                                                                                                                                                                                Penn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                      New
Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
```

3;

4;

Gaps

7

Indels

Length 240;

DB 22;

Score 135.5; DB 22; Pred. No. 2.1e-06; 5; Mismatches 25;

5

24.9%; ilarity 40.4%; Conservative

Similarity 36; Conser

Query Match

Local

Best

Matches

EKLFKEVKLE-

~

δ

50

'ACQN

-ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAV

ESLFREIRAQARALRAAAEPPPDPGPGAGSRKANLGGRRRORTALAGTRGXXGS-

79

YRQFCRGGGGGGGGGGGGGGGGGMSK

51

Q

53

d

Search completed: October 24, 2002, 09:27:58 Job time : 41.6861 secs

	·		•		
		•			
		Ç.,			
					•

---

```
updates/sec
                                                                                                                        Seconds
                                                                                                                                                                                                                                             ....DDFDLDVVKVSKQDSKITPQ 100
                                                                                                                     ; Search time 152.555 (without alignments) 230.724 Million cell
                                                                                                                                                                                                                                                                                                                                                                                       3502263
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/USO6_COMB.pep:*
2: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
4: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
5: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
6: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
7: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
8: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
10: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
11: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
12: /cgn2_6/ptodata/1/paa/USO8_COMB.pep:*
13: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
14: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
15: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
16: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
17: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
18: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
19: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
24: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
25: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
26: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
27: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
28: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
29: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
20: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
21: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
22: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
23: /cgn2_6/ptodata/1/paa/USO9_COMB.pep:*
 5.1.3
Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                         satisfying chosen parameters
                                                                                                                                                                                                                                                                                                                                               351980561 residues
  version
- 2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      summaries
                                                                                  model
                                                                                                                      2002, 09:28:06
                                                                                                                                                                                                 US-09-894-030-3
545
1 MEKLFKEVKLEELENOKGSG...
                                                                                                                                                                                                                                                                                                          0.5
                                                                                  - protein search, using sw
GenCore (c) 1993
                                                                                                                                                                                                                                                                                                          Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 s
                                                                                                                                                                                                                                                                                                                                                                                                                             seq length: 0
seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                 seds,
                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 ,
                     Copyright
                                                                                                                      October 24,
                                                                                                                                                                                                                                                                                                                                               3502263
                                                                                                                                                                                                                                                                                                                                                                                       hits
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of
                                                                                                                                                                                                                              score:
                                                                                                                                                                                                                                                                                     table:
                                                                                                                                                                                                                                                                                                                                                                                       number
                                                                                                                                                                                                                                                                                                                                                                                                                                 80
08
08
                                                                               OM protein
                                                                                                                                                                                                                                                Sequence:
                                                                                                                                                                                                                                                                                                                                                Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Database
                                                                                                                                                                                                     Title:
Perfect
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Maximum
                                                                                                                                                                                                                                                                                      Scoring
                                                                                                                                                                                                                                                                                                                                                                                                                                 Minimum
                                                                                                                        Run on:
```

to have a being printed cted by chance to of the result being score distribution results predicted by l to the score of of the total scor score greater than or equal and is derived by analysis o of the number is Pred. No.

/cgn2\_6/ptodata/1/paa/US60\_COMB.pep:\*

			Description	iland & opnomboo	trade of someon	Sequence 3, Appli	Sequence 3, Appli	Sequence 5, Appli		Sequence 5, Appli	Sequence 7, Appli
SUMMAKIES			ID		0 0 0 0 0 0 0 0	US-09-893-600-3	US-09-894-030-3	PCT-US98-14547-5	PCT-US98-14547-5	US-09-462-478A-5	PCT-US98-14547-7
			DB		1	22	22	-1	~	18	7
			Length DB	100	2	100	100	56	56	26	37
,	ďΡ	Query	4	10001		100.0	100.0	55.8	55.8	55.8	38.7
			Score	6 1 1	7	545	545	304	304	304	211
		Result	No.	 	+	7	m	4	ហ	9	7

Sequence 7, App Sequence 7, App Sequence 18, Ap Sequence 18, Ap Sequence 18, Ap Sequence 18, Ap Sequence 56, Ap Sequence 56, Ap	2220 Sequence 22220  1 Sequence 11, App  Sequence 2, App  Sequence 895, Sequence 19, App  Sequence 19,	3504 3504, 364ence 10485, 364ence 10484, 18 364ence 118, Ap 37 36quence 197, Ap 36 36 36 36 36 36 36 36 36 36	2221 Sequence 22221, 2219 Sequence 22219, 0 Sequence 30, Appl Sequence 30, Appl Sequence 26550, A 3599 Sequence 33599, 8971 Sequence 9, Appl Sequence 9, Appl Sequence 9, Appl	6456 Sequence 1645 5745 Sequence 2574 3603 Sequence 4360 222 Sequence 9222 222 Sequence 9222 Sequence 5, Ap
PCT-US98-14547- B US-09-462-478A PCT-US98-14547- PCT-US98-14547- B US-09-462-478A US-08-085-126-5 US-08-438-114-5	B US-09-417-507- PCT-US99-05250- 9 US-09-515-886- 9 US-09-538-092- 8 US-09-497-822-	uS-09-41/-50/ 1 US-09-708-427 1 US-09-708-427 PCT-US00-29360 PCT-US01-11988 PCT-US02-08123 PCT-US02-08277 PCT-US02-08277	8 US-09-417-507 8 US-09-417-507 US-08-333-576A US-08-362-670- PCT-US01-00663 2 US-09-864-761 6 US-60-236-359 7 US-09-394-272	6 US-09-248-796-6 US-60-096-409-8 US-09-417-507-1 US-09-733-089-2 US-09-816-660-PCT-US01-00358-PCT-US99-17885-
37 37 37 37 96	დდთთოო	/ 0 2 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9	285 188 188 240 7 240 7 101 101 101 101 208 1	001000000000000000000000000000000000000
6677788			22222222222222222222222222222222222222	
1100044	-t · · · · ·	# <b># # # # #</b>	139.55 139.55 135.55 134.55 134.55	25, 23, 23, 23, 23, 23, 23, 23, 23, 23, 23
8 10 11 12 13	15 16 18 19 20	0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	, w w w w w w w w w w w w w w w w w w w	0 6 4 4 4 4 4 4 6 0 0 0 0 0 0 0 0 0 0 0

ALIGNMENTS

RESULT 1
US-09-893-499-3
; Sequence 3, Application US/09893499
; GENERAL INFORMATION:
; APPLICANT: Hansen, J. Norman
; TITLE OF INVENTION: Construction and Screening of Lantibody Display Libraries
; FILE REFERENCE: 108172-00056
; CURRENT APPLICATION NUMBER: US/09/893,499
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,449
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1 ö MEKLFKEVKLEELENOKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGG 60 Gaps ö Length 100; Indels ; OTHER INFORMATION: The peptide sequence of sunA-PG20-SL. US-09-893-499-3 Score 545; DB 22; Pred. No. 1.4e-49; Mismatches 0; Query Match
Best Local Similarity 100.0%; I
Matches 100; Conservative 0; ORGANISM: Artificial Sequence FEATURE: SEQ ID NO 3 LENGTH: 100 PRT Н Н ga δy

GGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKITPQ 100

ŏ

SUBTILIS

PRODUCED BY BACILLUS

```
PCT-US98-14547-5

Sequence 5, Application PC/TUS9814547B

GENERAL INFORMATION:

APPLICANT: UNIVERSITY OF MARYLAND

TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PROD

FILE REFERENCE: 8172-8068

CURRENT APPLICATION NUMBER: PCT/US98/14547B

CURRENT FILING DATE: 1998-07-17

EARLIER FILING DATE: 1997-07-18

NUMBER OF SEQ ID NOS: 18

SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                             Pre-Sublancin 168
                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                 subtilis 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           subtilis 168
                                                                                                                                                                                                                                                                                                                                                         55.8%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    55.8%;
llarity 100.0%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2.0
                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PatentIn Ver.
                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 56; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     . Similarity
56; Conser
                                                                                                                                                                                                                                                                TYPE: PRT
ORGANISM: Bacillus
                                                                                                                                                                                                                                                                                            ; FEATURE:
; OTHER INFORMATION:
PCT-US98-14547-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
ORGANISM: Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 6
US-09-462-478A-5
                                                                                                                                                                                                                         SEQ ID NO 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Pa
SEQ ID NO 5
LENGTH: 56
                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Dp
                                                                                                                                                                                                                                                                                                                                                                                                                        ŏλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Sequence 3, Application US/09894030

Sequence 3, Application US/09894030

GENERAL INFORMATION:
APPLICANT: Hansen, J. Norman

TITLE OF INVENTION: Construction of a Structural Variant of Sublancin to Facilitate i TITLE OF INVENTION: Isolation and Use in Bioremediation of Environmental Contamination TITLE OF INVENTION: positive Spore Formers such as Bacillus anthrasis
FILE REFERENCE: 108172-00058

CURRENT APPLICATION NUMBER: US/09/894,030

CURRENT FILING DATE: 2001-06-29

PRIOR FILING DATE: 2000-06-29

PRIOR FILING DATE: 2000-06-29
                                                                                                                                         that Displays
                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                111111
CRGGGG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                RGGGG 60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         .
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           100;
                                                                                                                  APPLICANT: Hansen, J. Norman
TITLE OF INVENTION: Construction of a Strain of Bacillus subti-
TITLE OF INVENTION: Sublancin Lantibiotic on the Surface of t.
FILE REFERENCE: 108172-00057
CURRENT APPLICATION NUMBER: US/09/893,600
CURRENT FILING DATE: 2001-06-29
PRIOR FILING DATE: 2000-06-29
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Tag
                                                                                                                                                                                                                                                                                                                                                                                                       Length 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sublancin-His
                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                       OTHER INFORMATION: The peptide sequence of sunA-PG20-SL. -09-893-600-3
61 GGGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKITPQ 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; DB 22;
1.4e-49;
hes 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       100.0%; Score 545; DB 22;
100.0%; Pred. No. 1.4e-49;
tive 0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKITPQ
                                                                                                                                                                                                                                                                                                                                                                                                    Score 545; D
Pred. No. 1.4
; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                sequence for
                                                                                      Sequence 3, Application US/09893600 GENERAL INFORMATION:
APPLICANT: Hansen, J. Norman
                                                                                                                                                                                                                                                                                                                                                                                                  Match 100.0%; S Local Similarity 100.0%; F nes 100; Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             INFORMATION: The peptide
                                                                                                                                                                                                                                                                                                                       ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
ORGANISM: Artificial Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn version 3.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 Similarity 100; Conser
                                                                                                                                                                                                                                                                    ID NO 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 3
                                                          RESULT 2
US-09-893-600-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -09-894-030-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-09-894-030-3
                                                                                                                                                                                                                                                                                                           PRT
                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Loca1
                                                                                                                                                                                                                                                                                                                                       FEATURE:
                                                                                                                                                                                                                                                                                             CENGTH:
             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      61
                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                            TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OTHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              61
                                                                                                                                                                                                                                                                                                                                                                                                                     Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δŽ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ά
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Q
```

```
SUBLANCIN LANTIBLOTIC PRODUCED BY BACILLUS SUBTILIS 168
                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                                                                                                      168
                                                                                                                                                                                                                                                                         APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: PCT/US98/14547C
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: US/60/053,035
FARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                     56
                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
 Length 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                              Indels
ore 304; DB 1; 1 red. No. 2.2e-24; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 304; DB 1; Pred. No. 2.2e-24; D; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: US/09/462,478A
CURRENT FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: US/60/053,035
PRIOR FILING DATE: 1997-07-18
 Score 304;
Pred. No.
                                                                                                                                                                                            RESULT 5
PCT-US98-14547-5
Sequence 5, Application PC/TUS9814547C
GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; OTHER INFORMATION: Pre-Sublancin 168 PCT-US98-14547-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Sequence 5, Application US/09462478A GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
```

d d

δy

```
US-09-462-478A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID NO 18
LENGTH: 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                              ID NO 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
                                                                                                                                                                                                                                                                                                                    FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGIH:
                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                         a
                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                               ò
               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ö
                                                                                                                                                                ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 7, Application PC/TUS9814547C
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS 168
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: PCT/US98/14547C
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: US/60/053,035
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                            APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS
FILE REFERENCE: 8172-8068
                                                                                                                                     Length 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; Le
9.3e-15;
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 211; DB 1;
Pred. No. 9.3e-15;
0; Mismatches 0;
                                                                                                                                    Score 304; DB 18;
Pred. No. 2.2e-24;
0; Mismatches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                       CURRENT APPLICATION NUMBER: PCT/US98/14547B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: US/60/053,035
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 7
LENGTH: 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 211;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     No
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Pred.
                                                                                                                                                                                                                                                                         PCT-US98-14547-7; Sequence 7, Application PC/TUS9814547B; GENERAL INFORMATION:
                                                                                             ; OTHER INFORMATION: Pre-Sublancin 168 US-09-462-478A-5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; OTHER INFORMATION: Prosublancin 168
PCT-US98-14547-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OTHER INFORMATION: Prosublancin 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           38.7%; Scu
100.0%; Pi
0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                38.7%; Sc.
100.0%; P:
- 0;
                                                                                                                                    55.8%; Sc. larity 100.0%; F. Conservative 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ORGANISM: Bacillus subtilis 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     subtilis 168
                                                                   168
                                                                  subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity 100
37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
            Patentin Ver.
NUMBER OF SEQ ID NOS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 37; Conser
                                                                                                                                                  Similarity
56; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Bacillus
                                                    TYPE: PRT ORGANISM: Bacillus FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PCT-US98-14547-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PCT-US98-14547-7
                          SEQ ID NO 5
LENGTH: 56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TON GI
                                                                                                                                        Query Match
                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FEATURE:
               SOFTWARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQ ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 8
                                                                                                                                                     Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
                                                                                                                                                                                                                                                               RESULT
```

q

ð

```
NN: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS SUBTILIS 168
8172-8072
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SUBTILIS 168
                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS
FILE REFERENCE: 8172-8068
CURRENT APPLICATION NUMBER: PCT/US98/14547B
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: US/60/053,035
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
                                                                                                                                                                                                                                                                                                                                                                                              Length 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a dehydrogenated Ser (Dha)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a dehydrogenated Thr (Dhb)
                                                                                                                                                                                                                                                                                                                                                                                                                                      ÷
                                                                                                                                                                                                                                                                                                                                                                                               Score 211; DB 18;
Pred. No. 9.3e-15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20 GLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR 56
Sequence 7, Application US/09462478A; GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND; TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PF; FILE REFERENCE: 8172-8072; CURRENT APPLICATION NUMBER: US/09/462,478A; CURRENT FILING DATE: 2000-04-17; PRIOR APPLICATION NUMBER: US/60/053,035; PRIOR FILING DATE: 1997-07-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1 GLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 18, Application PC/TUS9814547B GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                    Prosublancin 168
                                                                                                                                                                                                                                                                                                                                                                                38.7%; SCV
100.0%; Pr
                                                                                                                                                                                                                                              LENGTH: 37
TYPE: PRT
ORGANISM: Bacillus subtilis 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ORGANISM: Bacillus subtilis 168
                                                                                                                                                                                                             2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LOCATION: (19)
OTHER INFORMATION: Xaa is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        OTHER INFORMATION: Xaa is
                                                                                                                                                                                          18
                                                                                                                                                                                      NUMBER OF SEQ ID NOS: 1:
SOFTWARE: Patentin Ver.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID (14)..(29)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                тн гоетн
(19)..(22)
                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
Matches 37; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY: DISULFID LOCATION: (7)..(36)
                                                                                                                                                                                                                                                                                                                                    ; OTHER INFORMATION:
US-09-462-478A-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NAME/KEY: MOD_RESTOCATION: (16)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NAME/KEY: MOD_RES
LOCATION: (19)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PCT-US98-14547-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -US98-14547-18
```

Length 37;

ï

B

Score 202;

37.1%;

ó

.rapm

us-09-894-030-3

Page

```
OTHER INFORMATION: Xaa
NAME/KEY: THIOETH
LOCATION: (19)..(22)
                        NAME/KEY: DISULFID LOCATION: (7)..(36) NAME/KEY: MOD_RES LOCATION: (16)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Washington
                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 35; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ; TOPOLOGY: 11;
; MOLECULE TYPE:
US-08-085-126-56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          20004
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT 14
US-08-438-114-56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STREET:
CITY: Wa
STATE: D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
                                                    ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SUBTILIS 168
                                                                                                                                                                                                                                                                                                                                                                        SUBTILIS
                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                    ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 S
                                                                                                                                                                                                                                                                                                                              APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: PCT/US98/14547C
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: US/60/053,035
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 18
LENGTH: 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 18, Application US/09462478A
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MARYLAND
TITLE OF INVENTION: SUBLANCIN LANTIBIOTIC PRODUCED BY BACILLUS
FILE REFERENCE: 8172-8072
CURRENT APPLICATION NUMBER: US/09/462,478A
CURRENT FILING DATE: 2000-04-17
PRIOR APPLICATION NUMBER: US/60/053,035
PRIOR FILING DATE: 1997-07-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 37;
                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (Dha)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (qya)
8.4e-14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .. ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 202; DB 1;
Pred. No. 8.4e-14
); Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   h NAME/KEY: MOD_RES
LOCATION: (19)
color="block" colo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INFORMATION: Xaa is a dehydrogenated Ser
                                                                                             56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 GLGKAQCAALWLQCAXGGXIGCGGGAVACQNYRQFCR 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           20 GLGKAQCAALWLQCASGGTIGCGGGAVACQNYROFCR
                                   Mismatches
                           .
So.
                                                                                                                                                                                                                                                     PCT-US98-14547-18; Sequence 18, Application PC/TUS9814547C; GENERAL INFORMATION:
                        Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE: PRT
ORGANISM: Bacillus subtilis 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  37.1%;
ilarity 94.6%;
Conservative
                      94.68;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE: PRT
ORGANISM: Bacillus subtilis
                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NUMBER OF SEQ ID NOS: 1
SOFTWARE: PatentIn Ver.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DISULFID (14)..(29)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (19)..(22)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 NAME/KEY: DISULFID LOCATION: (14)..(29)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DISULFID (7)..(36)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity
Matches 35; Conser
                     Similarity 35; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MOD_RES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           THIOETH
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-09-462-478A-18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQ ID NO 18
LENGTH: 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FEATURE:
NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        LOCATION:
                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FEATURE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                                                                  20
                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OTHER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQ ID
                                                                                                                                                                                                                                    RESULT
                                                                                                                                                      d
                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               a
```

```
ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                  ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            0;
                                                                                                                    Length 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 396;
                                                                                                                                                                                                                                                                                                                 Robert C.
Larick E.
CHELYDRYZYMES AND NEPAZYMES, AND USES
THEREFOR
236
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (Dha)
                                                                         a dehydrogenated Thr (Dhb)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.25
                                                                                                                   Score 202; DB 18;
Pred. No. 8.4e-14;
); Mismatches 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 144; DB 4;
Pred. No. 1.6e-06;
Mismatches 17
dehydrogenated Ser
                                                                                                                                                                            28,005
IR: LADNER=10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin Release #1.0, or current APPLICATION DATA:
APPLICATION NUMBER: US/08/085,126
FILING DATE: 25-JUN-1993
CLASSIFICATION: 435
                                                                                                                                                                                                                                                     RESULT 13
US-08-085-126-56; Application US/08085126; Sequence 56, Application US/08085126; GENERAL INFORMATION:
APPLICANT: LADNER, Robert C.
APPLICANT: CANNON, Larick E.
TITLE OF INVENTION: CHELYDRYZYMES;
TITLE OF INVENTION: THEREFOR
                                                                                                                                                                                                                                                                                                                                                                                                                         Z
                                                                                                                                                                                                                                                                                                                                                                                                           3: BROWDY AND NEIMARK 419 Seventh Street, N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9;
                                                                                                                                                  ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
                                                                                                                  37.1%;
illarity 94.6%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26.4%;
illarity 53.1%;
Conservative (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ATTORNEY/AGENT INFORMATION:
  ಡ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  NAME: COOPER, Iver P. REGISTRATION NUMBER: 28 REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         396 amino acids
amino acid
iY: linear
                         LOCATION: (19)..(22)
NAME/KEY: MOD_RES
LOCATION: (19)
OTHER INFORMATION: Xaa is
09-462-478A-18
15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy d:
COMPUTER: IBM PC COMP?
OPERATING SYSTEM: PC-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                         NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRESS
ADDRESSEE: BROWDY AI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
```

```
APPLICANT: KEITH G. WEINSTOCK ET AL.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ASPERGILLUS
TITLE OF INVENTION: FUMIGATUS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: PATH99-10
CURRENT APPLICATION NUMBER: US/09/417,507
CURRENT FILING DATE: 1999-10-14
NUMBER OF SEQ ID NOS: 44312
SEQ ID NO 22220
LENGTH: 276
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 396
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 27
                                                                          AND USES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                     CITY: Washing
CITY: Washing
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/438,114
FILING DATE: 08-MAY-1995
FILING DATE: 08-MAY-1995
                                                        Larick E.
CHELYDRYZYMES AND NEPAZYMES,
THEREFOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              )6;
17;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 144; DB 8;
Pred. No. 1.6e-06;
Mismatches 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 143;
                                                                                                                                                                                                                                                                                                                                                                FILING DATE: 08-MAY-1>>
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/308,793
FILING DATE: 19-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/085,126
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, Iver P.
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: LADNER-10F
TELECOMMUNICATION INFORMATION:
TELEFAX: 202-628-5197
TELEFAX: 202-737-3528
TELEFAX: 202-737-3528
TELEFAX: 396 amino acids
"VOR: amino acid
"'VOR: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LADNER-10B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SSULT 15
S-09-417-507-22220
Sequence 22220, Application US/09417507
GENERAL INFORMATION:
                                                                                                                                                   E: BROWDY AND NEIMARK
419 Seventh Street, N.W.,
Sequence 56, Application US/08438114;
GENERAL INFORMATION:
APPLICANT: LADNER, Richard C.
APPLICANT: CANNON, Larick E.
TITLE OF INVENTION: CHELYDRYZYMES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 240
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Ouery Match
Best Local Similarity 53.1%;
Matches 26; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     26.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; TYPE: PRT; ORGANISM: A.fumigatus
US-09-417-507-22220
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRANDEDNESS: Si

TOPOLOGY: linear

MOLECULE TYPE: pro

US-08-438-114-56
                                                                                                                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

•.
•

Appli Appli

Sequence Sequence Sequence

Sequence Sequence Sequence

ALIGNMENTS

777723344

Sequence Sequence Sequence Sequence Sequence Sequence

Sequence Sequence

.rai

 $\boldsymbol{\alpha}$ ı

```
US-08-220-033-9
US-08-465-491-2
US-08-465-491-9
US-08-986-617-2
US-08-986-617-7
US-08-986-617-9
US-08-986-617-9
US-08-977-767-3
US-08-900-230-3
US-08-960-190A-25
US-08-960-190A-25
US-08-960-190A-25
US-08-960-190A-25
US-08-981-254-7
US-08-981-270A-7
US-09-013-587-7
US-09-013-587-7
US-09-013-587-7
US-09-013-587-7
                                                                                                                                                                                                                         Sequence 11, Application US/09041886 Patent No. 6235872 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -CRGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ω
..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             26.1%;
ilarity 42.5%;
Conservative
    4000000000044500
  protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local Similarity
nes 37; Conser
  TOPOLOGY: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; MOLECULE TYPE:
US-09-041-886-11
                                                                                                                                                                                                          RESULT 1
US-09-041-886-11
   Query Match
Best Local S
Matches 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      55
   Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3, Appli
15, Appli
2, Appli
11, Appl
26, Appl
2, Appli
7, Appli
9, Appli
2, Appli
7, Appli
7, Appli
7, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                             1, Appli
3, Appli
9, Appli
6, Appli
7, Appli
7, Appli
3, Appli
3, Appli
3, Appli
3, Appli
15, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                    APP1
APP1
APP1
APP1
                                                                                 updates/sec
                                                                                                                                                                                                                                                                                                                                    cted by chance to have a of the result being printed score distribution.
                                                                                                                                                                                                                                                                                                                                                                                                                 111,
30,
30,
30,
                                                                Seconds
                                                                                                                    .. DDFDLDVVKVSKQDSKITPQ 100
                                                                                                                                                                                                                                                                                                                                                                                                 Description
                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
Seq
                                                                S
                                                               ; Search time 15.328 (without alignments) 159.348 Million cell
                                                                                                                                                                                                                                                              Issued_Patents_AA:*
    /cgn2_6/ptodata/1/laa/5A_COMB.pep:*
    /cgn2_6/ptodata/1/laa/5B_COMB.pep:*
    /cgn2_6/ptodata/1/laa/6A_COMB.pep:*
    /cgn2_6/ptodata/1/laa/6B_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*
    /cgn2_6/ptodata/1/laa/PCTUS_COMB.pep:*

                      Ltd
           5.1.3
Compugen
                                                                                                                                                                                                                                                                                                                                                                                                               US-09-041-886-11
US-08-352-670B-30
US-08-333-576C-30
US-08-808-324-30
PCT-US94-14030A-30
US-09-052-995-1
US-09-053-003-40
US-08-864-038A-3
US-08-8157-349-9
US-08-8157-349-9
US-08-8157-349-9
US-08-8157-349-9
US-08-8157-349-9
US-08-581-528A-6
US-08-64-07799-6
US-08-486-421-3
US-08-486-421-3
US-08-486-809-3
US-08-486-809-3
US-08-486-809-3
US-08-207-904-15
US-08-207-904-15
US-08-986-617-26
US-08-986-617-26
US-08-986-617-26
US-08-981-525-7
US-08-220-033-2
                                                                                                                                                                                                                                                                                                                                   number of results predicted a or equal to the score of tanalysis of the total score
                                                                                                                                                                                 chosen parameters
                                                                                                                                                              residues
                                                                                                                                                                                                                                                                                                                                                                        SUMMARIES
             version
- 2002
                                                                                                                                                                                                                          Post-processing: Minimum Match 0%
.Maximum Match 100%
Listing first 45 summaries
                                               model
                                                               09:27:01
                                                                                                  US-09-894-030-3
545
1 MEKLFKEVKLEELENQKGSG
                                               SK
                                                                                                                                              0
                                                                                                                                                               24425594
            GenCore
(c) 1993
                                                                                                                                              Gapext
                                                using
                                                                                                                                                                                                 2000000000
                                                               2002,
                                                                                                                                                                                                                                                                                                                                                                                                                    4164546060544140444001444
                                                                                                                                                                                 satisfying
                                                                                                                                                                                                                                                                                                                                                                                                  No. is the number greater than or eq i derived by analys
                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                seds,
                                               protein search,
                      Copyright
                                                                                                                                    BLOSUM62
Gapop 10.0
                                                                24,
                                                                October
                                                                                                                                                                                                 length:
length:
                                                                                                                                                              231628
                                                                                                                                                                                                                                                                                                                                                                                      Query
Match
                                                                                                                                                                                                                                                                                                                                                                                                                  hits
                                                                                                                                                                                                                                                                       65.5.5.5.5
                                                                                                                                                                                of
                                                                                                                                                                                                                                                                                                                                                                                                           sed
sed
                                                                                                                                                                                                                                                                                                                                                                                                  core
                                                                                                                                                                                                                                                                                                                                     о<u>Р</u>
.
                                                                                                             score:
                                                                                                                                     table:
                                                                                                                                                                                                                                                                                                                                    Pred. N
score g
and is
                                                                                                                                                                                 number
                                                                                                                                                                                                 DB
DB
                                                OM protein
                                                                                                                                                                                                                                                                                                                                                                                                  S
                                                                                                                     Sequence:
                                                                                                                                                               Searched:
                                                                                                                                                                                                 Minimum
Maximum
                                                                                                                                     Scoring
                                                                Run on:
                                                                                                                                                                                                                                                                                                                                                                                        Result
No.
                                                                                                                                                                                                                                                                                                                                                                                                                   Title
                                                                                                                                                                                 Total
```

Peptides, Dependence and Methods of Use

4

S

918

Length

Score 142.5; DB 4; Pred. No. 5.2e-07; 3; Mismatches 22;

438

54

-AVACQNYRQF-

-GTIGCGGG-

```
NUMBER OF SEQUENCES: 37 CORRESPONDENCE ADDRESS:
                                                                                                           COMPUTER READABLE MEDIUM TYPE: FI
                                                                                                                                                                                                                                                                                                                                        INFORMATION FOR SEQ
                                                                                                                                                                                                                                                                                                                                                                                                         ; MOLECULE TYPE:
US-08-333-576C-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   02140
                                                                                               02140
                                        STREET:
CITY: Ca
STATE: M
COUNTRY:
ZIP: 021
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           αq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                   APPLICANT: Rosen, Vicki A.
APPLICANT: Wolfman, Neil
APPLICANT: Thomsen, Gerald H.
APPLICANT: Melton, Douglas A.
TITLE OF INVENTION: TENDON-INDUCING COMPOSITIONS NUMBER OF SEQUENCES: 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Douglas A.
TENDON-INDUCING COMPOSITIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 135.5; DB 1;
Pred. No. 7.2e-07;
5; Mismatches 25;
                                                                                                                                                                                                                                                                                                                           ZIP: UZ14U
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,670B
              465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202-D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        51 YRQFCRGGGGGGGGGGGGGGGGGGGMSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence 30, Application US/08333576C
Patent No. 6027919
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J.
                                                                                Sequence 30, Application US/08362670B atent No. 5658882 GENERAL INFORMATION: APPLICANT: Celeste, Anthony J.
                                                                                                                                                                                                                                       CORRESPONDENCE ADDRESS:
ADDRESSEE: GENETICS INSTITUTE,
STREET: 87 CambridgePark Drive
439 QLYGPCGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5;
                                                                                                                                                Rosen, Vica.
Wolfman, Neil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            32,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REFERENCE/DOCKET NUMBER: 52
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Rosen, Vicki A.
APPLICANT: Wolfman, Neil
APPLICANT: Thomsen, Gerald H
APPLICANT: Melton, Douglas A
TITLE ÖF INVENTION: TENDON-II
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARACTERISTICS: 240 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ..
0N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity 40.436; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Wozney, John
Rosen, Vicki
                                                                                                                                         Wozney, John
Rosen, Vicki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                               Cambridge
Massachusetts
                                                                                                                                                                                                                                                                                                                                                                                                                                                    514
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       linear
                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: DE CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EKLFKEVKLE-
                                                                                                                                                                                                                                                                                                               \mathbf{USA}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; MOLECULE TYPE:
US-08-362-670B-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               INFORMATION FOR
                                                                                                                                                                                                                                                                                                           COUNTRY: U
                                                     RESULT 2
US-08-362-670B-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US-08-333-576C-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                          APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local
                                                                                 Sequence 3(
Patent No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                              ADABLE FORM:
(PE: Floppy disk: IBM PC compatible; SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Rosen, Vic.
Wolfman, Neil
Thomsen, Gerald H.
Melton, Douglas A.
Melton, TENDON-INDUCING COMPOSITIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,324
FILING DATE: Herewith
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 135.5; DB 3;
Pred. No. 7.2e-07;
; Mismatches 25;
                                                                                                                                                 OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/333,576C
FILING DATE: No. 6027919ember 2, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             79
                                                                                                                                                                                                                                                                                  32,618
FR: 5202-B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             51 YRQFCRGGGGGGGGGGGGGGGGGGMSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --GGGGGGGGGGGGGGGGGGGAGR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 30, Application US/08808324
Patent No. 6284872
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J.
ADDRESSEE: GENETICS INSTITUTE, STREET: 87 CambridgePark Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ADDRESSEE: GENETICS INSTITUTE,
STREET: 87 CambridgePark Drive
                                                                                                                                                                                                                                                                                                     REFERENCE/DOCKET NUMBER: 52
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         24.9%;
Local Similarity 40.4%;
nes 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Wozney, July
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: Rosen, Vicki
APPLICANT: Wolfman, Neil
APPLICANT: Thomsen, Gera
APPLICANT: Melton, Dougl
TITLE OF INVENTION: TENI
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                         ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                                   Massachusetts
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Massachusetts
USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STREET: 87 Cambraction CITY: Cambridge STATE: Massachuse COUNTRY: USA
                                   Cambridge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             EKLFKEVKLE---
```

4;

```
5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Parent, Annette S.
REGISTRATION NUMBER: 42,058
REFERENCE/DOCKET NUMBER: 016
TELECOMMUNICATION INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24.3%;
illarity 63.4%;
Conservative (
                              24.9%;
Similarity 40.4%;
36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           201 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Pengguang
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CURRENT APPLICATION DATA APPLICATION NUMBER: U
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            peptide
                                                                                            EKLFKEVKLE----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RY: USA
94111-3834
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
Matches 26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 3: CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    99 GGPGGGGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Wu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ; TOPOLOGY: 11
; MOLECULE TYPE:
US-09-052-995-1
  PCT-US94-14030A-30
                                                                                                                                                                                                                                                                                                             APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
                                                                                                                                                                                                                                                    US-09-052-995-1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGIH:
                                Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loc
Matches
                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36
                                                                                            N
                                                                                                                                                                                      107
                                                                                                                                                                                                                                    RESULT
                                                                                                                        qq
                                                                                                                                                                                      Db
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Dp
                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δ
                                                                                                                                                                                                                                                                                4;
                                                                                                                                                                                                                                                                                                                                             106
                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                            EKLFKEVKLE------ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN
                                                                                                                                                                                                                                                                                23;
                                                                                                                                                                                                                                                    240;
                                                                                                                                                                                                                                                                                                                                         53 ESLFREIRAQARALRAAAEPPDDPGPGAGSRKANLGGRRRORTALAGTRGXXG
                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HARVARD COLLEGE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICANT: GENETICS INSTITUTE, INC.
APPLICANT: PRESIDENT AND FELLOWS OF HARVARD COLLITILE OF INVENTION: TENDON-INDUCING COMPOSITIONS NUMBER OF SEQUENCES: 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Version #1.25
                                                                                                                                                                                                                                               Score 135.5; DB 4;
Pred. No. 7.2e-07;
5; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/164,103
FILING DATE: 07-DEC-1993
APPLICATION NUMBER: US 08/217,780
FILING DATE: 25-MAR-1994
APPLICATION NUMBER: US 08/333,576
FILING DATE: 02-N0V-1994
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER: PCT/US94/14030A FILING DATE: Herewith
                                                                                                                                                                                                                                                                                                                                                                                          -----GGGGGGGGGGGGGGGGGGGAGR 129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TITLE OF INVENTION: 35
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENETICS INSTITUTE, INC.
STREET: 87 CambridgePark Drive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5202D-PCT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Sequence 30, Application PC/TUS9414030A GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                         51 YRQFCRGGGGGGGGGGGGGGGGGMSK
                                                             5202-D
                                                                                                                                                                                                                                                                            5;
                                                                       TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
                                               32,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      32,618
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30:
                                                                                          TELEFAX: 617 8/0 30
INFORMATION FOR SEQ ID NO: 30
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acid
TVPE: amino acid
                                                                                                                                                                                                                                            24.9%;
40.4%;
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
                              NAME: Lazar, Steven R. REGISTRATION NUMBER: 32, REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: Lazar, Steven R. REGISTRATION NUMBER: 32, REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               240 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEFAX: 617 0.0
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                              Similarity 40.
                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Massachusetts
                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cambridge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: H
CLASSIFICATION:
                                                                                                                                                                                 TOPOLOGY: 11

MOLECULE TYPE:
US-08-808-324-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PCT-US94-14030A-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     02140
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MOLECULE
                                                                                                                                                                                                                                                Query Match
Best Local
                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                     107
                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                          8
                                                                                                                                                                                                                                                                                                                                                                                                     pp
                                                                                                                                                                                                                                                                                                                                                                        Óχ
                                                                                                                                                                                                                                                                                                             ð
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ij
                                                                                                                     106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                              --ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN
                                        23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11;
                                                                                                                                                                                                                                                                                                                                                       AArion.
Sivaraja, Mohanram
Sivaraja, Mohanram
Strulovici, Berta
Flores, Osvaldo A.
Flores, Osvaldo A.
VENTION: High Throughput In Vitro Screening Assay
VENTION: for Transcription Modulators
VENTION: 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     201;
  Length 240;
                                                                                                                     53 ESLFREIRAQARALRAAAEPPPDPGPGAGSRKANLGGRRRQRTALAGTRGXXGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -GGGGGGGGGGGGGGGGG 128
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                          TITLE OF INVENTION: High Throughput In Vitro Sci
TITLE OF INVENTION: for Transcription Modulators
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 132.5; DB 4;
Pred. No. 1.2e-06;
0; Mismatches 4;
Score 135.5; DB 5;
Pred. No. 7.2e-07;
; Mismatches 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   018781-000600US
                                                                                                                                                                                    51 YRQFCRGGGGGGGGGGGGGGGGGMSK 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  MBER: US/09/052,995
31-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 7
US-09-053-003-40
; Sequence 40, Application US/09053003
; Patent No. 6207391
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                      Sequence 1, Application US/09052995
Patent No. 6183956
GENERAL INFORMATION:
```

.rai

```
STATE: Mie-prefecture
COUNTRY: JAPAN
ZIP: 514-01
                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE TYPE: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Boston
                                                                                                                                                                                                                                                                                                                                                                                                                             amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ORGANISM: P
CELL TYPE:
FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE:
STREET: 13
CITY: Bost
STATE: MA
                                                                                                                                                                                                                                                                                                                                                                                                                                           TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -08-835-099A-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Kunio NAKASHIMA et al.

FUNTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
FENTION: CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
FENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
FENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
FENTION: TO SAID POLYPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    þe
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 20
                                     High-Throughput Screening Assays for Modulators of STAT4 and STAT6 Activity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             /note= "Gly at positions 105-201 may present or absent"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 may be
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         /product= "OTHER"
/note= "Gly at positions 1-97
present or absent"
APPLICANT: McKinney, Juan
TITLE OF INVENTION: High-Throughput Screeuing
TITLE OF INVENTION: Modulators of STAT4 and STAT6
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
ADDRESSEE: Townsend and Townsend and Crew LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 --GGGGGGGGGGGGGGGGG 128
                                                                                                                                                                                                                                                                            #1.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 132.5; DB 4;
Pred. No. 1.2e-06;
0; Mismatches 4;
                                                                                                                                                STATE.

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

"OMPUTER: IBM PC compatible

"OMPUTER: IBM PC compatible

"OMPUTER: TRM PC-DOS/MS-DOS

"TRM Release #1.0, Version #1
                                                                                                                                                                                                                                             COMPUTER: 1DF FC COMPUTER: Detact By Composition #1 SOFTWARE: Patentin Release #1.0, Version #1 CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/053,003
FILING DATE: 31-MAR-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Timothy L.
REGISTRATION NUMBER: 35,367
REFERENCE/DOCKET NUMBER: 018781-000800US
TELEPHONE: (415) 576-0200
TELEPHONE: (415) 576-0200
TELEPHONE: (415) 576-0200
TELEPHONE: CHARACTERISTICS:
LENGTH: 201 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                "OTHER"
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3, Application US/08864038A
). 6001592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               /product=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Match 24.3%;
Local Similarity 63.4%;
nes 26; Conservative
                                                                                                                       STREET: Two Embarcadero
CITY: San Francisco
STATE: California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Modified-site
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ADDRESS:
812-5 Hirano
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              peptide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            LOCATION: 1..97
OTHER INFORMATION:
OTHER INFORMATION:
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NUMBER OF SEQUENCES:
CORRESPONDENCE ADDRES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Isshinden
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LOCATION: 105..20
OTHER INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CTHER INFORMATION:
CTHER INFORMATION:
US-09-053-003-40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Patent No. 6001592
GENERAL INFORMATION:
APPLICANT: Kunio
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Tsu-city
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -- 5555555555 66
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOLECULE TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         STRANDEDNESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ADDRESSEE:
STREET: IS
CITY: TSU-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             NAME/KEY:
LOCATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME/KEY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-864-038A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FEATURE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               П
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ð
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           --GGSGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        738;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 THEIR PRODUCTION
                  1.44 MB storage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CUSMAN, LLP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 129.5; DB 3;
Pred. No. 9.2e-06;
3; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            COMPUTER: IBM Compatible OPERATING SYSTEM: DOS SOFTWARE: FastSEQ for Windows Version 2.0 CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       E (by experiment)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ď
                            COMPUTER: IBM Compatible
COMPUTER: IBM Compatible
OPERATING SYSTEM: Microsoft Windows
COFTWARE: Word Perfect 6.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ROBERTS
                                                                                                      APPLICATION NUMBER: US/08/864,038A FILING DATE: May 28, 1997
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                   TELECOMMUNICATION NUMBER: 22,389
REFERENCE/DOCKET NUMBER: F-5610
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)986-2340
TELEFAX: (212)953-7733
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: NISHI, Kaznori
APPLICANT: KAWAMOTO, TOMOhiro
TITLE OF INVENTION: NOVEL PROTEINS,
TITLE OF INVENTION: AND USE
NUMBER OF SEQUENCES: 18
                   3.50 inch,
                                                                                                                                                                                                                                                                                                                                                                                                                                             Pinctada fucata
mantle epithelial cell
                                                                                                                                                         APPLICATION NUMBER: JP 8-184459
FILING DATE: 15-July-1996
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- 6866866668-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 9, Application US/08835099A Patent No. 5874277
GENERAL INFORMATION:
APPLICANT: SHINTANI, Yasushi APPLICANT: KAWAMOTO, Tomohiro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3: DIKE, BRONSTEIN, 130 Water Street
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23.8%;
larity 41.0%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      to 738
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -GMSKFDDFDLD 86
                                                                                                                                                                                                                NAME: C. Bruce Hamburg REGISTRATION NUMBER: 2
                  Diskette,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NAME/KEY: peptide;
LOCATION: from 1 to 7:
IDENTIFICATION METHOD:
US-08-864-038A-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Diskette
                                                                                       CURRENT APPLICATION DATA
                                                                                                                                                                                                                                                                                                                                                                                                           protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
    FORM:
                                                                                                                                                                                                                                                                                                                                                                                         linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
COMPUTER READABLE MEDIUM TYPE: D:
```

S

Thu

```
linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        l Similarity
25; Conserv
                                                               STRANDEDNESS: si

TOPOLOGY: linear

MOLECULE TYPE: pro

US-09-157-349-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ALAGTRGAOGS --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CLASSIFICATION:
ATTORNEY/AGENT IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                   USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: 11

MOLECULE TYPE:
US-08-581-528A-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                               92037
                                                                                                                                                                                                                 9 KGGGGG--
                                                                                                                                                                                                                                                                                                                                                                                                         STREET:
CITY: La
STATE: C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
Matches 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12
                                                                                                                                                                                        17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ò
                                                                                                                                                                                         δŽ
                                                                                                                                                                                                                  Ωp
                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                                                                                                                                                           12;
                                                                                                                                                                                                                                                                                                                                          Length 268
                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 10
US-09-157-349-9
; Sequence 9, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kaznori
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                         Score 129; DB 2;
Pred. No. 3.8e-06;
0; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/157,349
APPLICATION NUMBER: US/08/835,099A FILING DATE: 04-APR-1997 CLASSIFICATION: 514
                              CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 083649/1996
FILING DATE: 05-APR-1996
APPLICATION NUMBER: 97105508.2
FILING DATE: 03-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Resnick, David S
REGISTRATION NUMBER: 34,235
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    08/835,099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICATION NUMBER: 97105508. FILING DATE: 03-APR-1997. ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
                                                                                                                                               TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
                                                                                                                                                                                                                    ..
                                                                                                                                                                                                                                                                                                                                       th 23.7%;
Similarity 50.0%;
30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         IBM Compatible
                                                                                                                                                                                                                               CHARACTERISTICS: 268 amino acids amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LE FORM:
Diskette
                                                                                                                                                                 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IOR APPLICATION DATA: APPLICATION NUMBER:
                                                                                                                                                                                                                                                                        single
                                                                                                                                                                                                                                                                                                  protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    NAME: Resnick, Davider: REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                     STRANDEDNESS: sir
TOPOLOGY: linear
MOLECULE TYPE: pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE MEDIUM TYPE: Di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                02109
                                                                                                                                                                                                                                                                                                                                                                                                                           9 KGGGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COMPUTER:
                                                                                                                                                                                                                                                                                                               US-08-835-099A-9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COUNTRY:
                                                                                                                                                                                         TELEFAX:
                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                            TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PRIOR
                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                ŏ
```

```
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                   KGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12;
                                                                                                                                                                                     12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 161;
                                                                                                                                                  Length 268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGGGGTT
                                                                                                                                                                                                                                                                                                     SSULT 11
3-08-581-528A-6
Sequence 6, Application US/08581528A
Patent No. 5986058
GENERAL INFORMATION:
APPLICANT: Lee, Se-Jin
APPLICANT: Huynh, Thanh
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-7
NUMBER OF SEQUENCES: 21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --6666666666666666666666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 125; DB 2; Le
                                                                                                                                   Score 129; DB 3; Le
Pred. No. 3.8e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Suite 1400
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 07265/081001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/581,528A
FILING DATE: 03-Sept-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           P.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MMBER: US 08/089,670
09-JUL-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                        CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, I
STREET: 4225 Executive Square,
CITY: La Jolla
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Lisa A. Haile, Ph.D.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 0726
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                                  0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  22.9%;
nilarity 54.3%;
Conservative
  6
                                                                                                                                                  Best Local Similarity 50.0%;
Matches 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      161 amino acids
amino acid
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TELEFAX: 619/678-5099 INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       protein
                                                                    single
                                                                                                     protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER READABLE FORM: MEDIUM TYPE: Floppy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
```

rai.

```
ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-014-943A-2
                                                                                                                                                                                                                                                         TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
TYPE: a
                                                                                                                                                                                                                                                                                                      ; CLONE:
US-09-032-523-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STREET:
CITY: N
                                                                                                                                                                 TELEX:
               PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      72
                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       òχ
                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12;
                                           APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 161;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   --GGGGGGGGGGGGGGGGGAGR 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 125; DB 5;
Pred. No. 5.8e-06;
1; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HUMAN PROTEINASE MOLECULES
                                                                                                   EE: Spensley Horn Jubas & Lubitz
1880 Century Park East, Suite 500
Jos Angeles
California
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   2.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Inc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version
CURRENT APPLICATION DATA:
                                                                                                                                                                   LIPE: Floppy disk
CUMPUTER: IBM PC COmpatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Ve
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07799
FILING DATE: 08-JUL-1994
CLASSIFICATION:
ATTORNEY/AGENT
                                                                                                                                                                                                                                                                            MBER: PCT/US94/07799
08-JUL-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Incyte Pharmaceuticals,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              US/09/032,523
                                                                                                                                                                                                                                                                                                                              ATTORNEY/AGENT INFORMATION:
NAME: TUMARKIN, LISA A., PH.D.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: FD-2348
          Sequence 6, Application PC/TUS9407799
GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            7, Application US/09032523
                                                                                                                                                                                                                                                                                                                                                                                        TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 455-5100
TELEFAX: (619) 455-5110
FORMATION FOR SEQ ID NO: 6:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Jennifer
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3174 Porter Drive
Palo Alto
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Bandman, Olga
Hillman, Jennife
Corley, Neil C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local Similarity 54.3%;
Matches 25; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     IBM Compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                    QUENCE CHARACTERISTICS:
LENGTH: 161 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Guegler, Karl
Baugh, Mariah
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                      NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COMPUTER READABLE FORM:
MEDIUM TYPE: Diskett
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TITLE OF INVENTION: HI
NUMBER OF SEQUENCES: CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION NUMBER:
FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 7, Application Patent No. 6232454 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17 ALAGTRGAQGS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           USA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                       INFORMATION FOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        94304
                                                                                                                                 Los
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             \mathcal{C}_{\mathbf{A}}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ADDRESSEE:
                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PCT-US94-07799-6
PCT-US94-07799-6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                                                                                                  STREET:
CITY: LC
STATE: C
COUNTRY:
ZIP: 900
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           US-09-032-523-7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MOLECULE
                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ò
```

```
--GAGGGGGGGGGG 49
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24;
                                                                                                                                                                                                                                                                                                                                                                                                                          Length 266;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: Johnson, Edward M.
APPLICANT: Bergemann, Andrew D.
TITLE OF INVENTION: Cloning And Expression Of PUR Protein NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                         Score 125; DB 4; Le
Pred. No. 9.6e-06;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --GEGGGGGGGNATGGTIS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin Release #1.0, Verence application DATA:
APPLICATION DATA:
APPLICATION DATA:
APPLICATION NUMBER: US/08/014,943A
FILING DATE: 02)FEB)1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 6923-033
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
                                                                                        NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0479 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEPHONE: 650-845-4166
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         US/08/014,943A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2, Application US/08014943A
                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEE: Pennie & Edmonds: 1155 Avenue of the New York
                                                                                                                                                                                                                                                                                                                                                                                                                        22.9%;
llarity 44.6%;
Conservative
                                                                                                                                                                                                                      SEQUENCE CHARACTERISTICS:
LENGTH: 266 amino acids
TYPE: amino acid
STRANDEDNESS: single
                                                       FILING DATE: ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   322 amino acids
amino acid
CLASSIFICATION:
NOR APPLICATION DATA:
APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                              IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 164403
                                                                                                                                                                                                                                                                                                               linear
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 29; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 2, Application Patent No. 5545551 GENERAL INFORMATION:
                                                                                                                                                                                                                        INFORMATION FOR SEQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GGGGG 76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11111 50 GGGGG 54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9 KGGGGGG-
```

3;

7

```
9
                                                                                  9
                                                                                  Gaps
                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -GGGGGGGGGGGGGGG 46
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     46;
                                                                                  46;
                                                        Length 322;
                                                                                                                                                                                                                                                          Sequence 3, Application US/08486421
Patent No. 5672479
GENERAL INFORMATION:
APPLICANT: Johnson, Edward M.
APPLICANT: Bergemann, Andrew D.
TITLE OF INVENTION: CLONING AND EXPRESSION OF PUR PROTEIN NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 322;
                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Version #1.30
                                                                        )5;
10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 124; DB 1; Pred. No. 1.5e-05; 6; Mismatches 10;
                                                                                                                                                                               47 GGGAPGGLQHETQELASKRVDIQNKRFYLDV----KQNAK 82
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         --MSKFDD-----FDLDVVKVSKQDSK 96
                                                     Score 124; DB 1;
Pred. No. 1.5e-05
; Mismatches 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -- MSKFDD-----FDLDVVKVSKQDSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              COMPUTER: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Ver
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,421
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/470,911
FILING DATE: 06-JUN-1995
ATTORNEY/AGENT INFORMATION:
REGISTRATION NUMBER: 6923-053
TELEPHONE: (212) 790-9090
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 322 amino acid
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                                                                                                                                                                                                                                                                                                                                           STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               22.8%;
larity 38.0%;
Conservative (
                                                     th 22.8%; Similarity 38.0%; 38, Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein
               protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         unknown
  linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity 38; Conser
                                                                                                                                                                                                                                                                                                                                                                                                    New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TIFE.
TOPOLOGY: uni
TOPOLOGY: li
                                                                                                                                     10 OGGAALG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               10 OGGAALG-
             ; MOLECULE TY
US-08-014-943A-2
                                                                                                                                                                73 GGGG--
                                                                                                                                                                                                                                   RESULT 15
US-08-486-421-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ; MOLECULE TO US-08-486-421-3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          73 GGGG-
                                                     Query Match
Best Local S
Matches 38
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
Matches 38
                                                                                                                                                                                                                                                                                                                                                                                                    CITY:
STATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    47
                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Ob
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                 ò
                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      φ
                                                                                                             ğ
```

Search completed: October 24, 2002, 09:30:12 Job time: 18.3285 secs

		3			
**					
	•				
			•		
					<b>¥</b> .
					•

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.
```

OM protein - protein search, using sw model

Seconds October 24, 2002, 09:26:16; Search time 18.9781 (without alignments) 506.317 Million cell Run on:

updates/sec

US-09-894-030-3 545 1 MEKLFKEVKLEELENQKGSG.....DDFDLDVVKVSKQDSKITPQ 100

BLOSUM62 Gapop 10.0 , Gapext 0.5 Title: Perfect score: Sequence: Scoring table:

283138 seqs, 96089334 residues Searched: 283138

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_71:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* 1 2 8 4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

	Description	blancin 168	ည္က (၂)	cal p	drogen re	ich c	cal prot	e-rich	e-phosphate	ic	ne-rich	hetical	ne-rich	hetical	lpain (EC	meotic prof	ycine-rich pro	pothetical prot	ycine-ric	ycine-rich cell	ycine-rich prot	pothetical prot	tic protein	cinding prote	o-1,4-beta-glu	morphogeneti	gshell protein	alpain (EC 3.4	brillarin [vali	ycine-rich prot
SUMMAKIES	QI	1278													CIHUL	$\sim$	^,		S34666									CIPGL	A38712	~
	DB			•	•	•	` '	` '	` '	•	• •	•	•	•																
	Length	25		_			_	$\overline{}$	~	~~	. ~	. ~	$\sim$	$\overline{}$	10	$\sim$	_	ന	_	$\boldsymbol{\alpha}$	CD	$\omega$	•		01	u /	ľ	266		_
d	Query Match	55.8		٠.					_:	_:	~	~.	~	~	ω.	ω.	w.	ω.	m.	ω.	m.	m.	ო	m,	w.	α.	α.	•	ď	Ö
	Score	304		. ^	-	_		13	34.		m	ന	~	$\sim$	N	N	N	(A	ľ	ιv	ľ	( 1	1.1	25	•	7			•	$\Xi$
	sult No.	- 1	7	m	4	2	9	7	æ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29

<u> भिन्तिक से से से स</u>	subtilin precursor acp-22 protein - y acp-22 protein - y calpain (EC 3.4.22
A45036 G96010 A42424 T31611 T29282 T29173 H84709 S09860 C84470 A38743 A35628	NIBSSA S16063 S32224 CIRBL
ааааааааанаа	11221
322 152 280 1585 136 388 302 316 481	256 199 199 266
222.3 222.5 222.6 222.5 222.3 222.3 222.3	222.2
123.5 123.5 123.5 122.5 121.5 121.5 121.5 121.5	20. 20. 20.
33 33 33 34 35 36 37 40 40	4 4 4 4 4 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5

## ALIGNMENTS

 RESULT 1 T12783 sublancin 168 precursor - Bacillus subtilis phage SPBc2 C; Species: Bacillus subtilis phage SPBc2 C; Species: Bacillus subtilis phage SPBc2 C; Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999 C; Accession: T12783; H69719 C; Accession: T12783; H69719 R; Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D. submitted to the EMBL Data Library, August 1997 A; Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 pro A; Reference number: Z17583
 A; Accession: T12783 A; Status: translated from GB/EMBL/DDBJ A; Status: translated from GB/EMBL/DDBJ A; Status: translated from GB/EMBL/DDBJ A; Molecule type: DNA A; Molecule type: DNA A; Residues: 1-56 < LAZ> A; Residues: 1-56 < LAZ> A; Cross-references: EMBL: AF020713; NID: g3025478; PID: g3025497; PIDN: Azevedo, V.; Ber A; Cross-references: EMBL: AF020713; NID: g3025478; PID: Alloni, G.; Azevedo, V.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Errington, J.; Fabret, C.; Ferrari,
A.; Ehrlich, S.D.; Emmerson, F.T.; Entlan, N.D.; Ellisycon, S.; Galizzi, A.; Gal Nature 390, 249-256, 1997  Nature 390, 249-256, 1997  A; Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galizzi, A.; Hullo, M.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino Koetter, P.; Koningstein, G.; Krogh, S.; Rumano, M.; Kurita, H.; Masuda, S.; Mau A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau Rieger, M.; Rivolta, C.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl Rieger, M.; Tamakoshi, A.; Tanakoshi, A.; Tanamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida A; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.; Tanakoshi, A.; Tanakoshi, B.; Joshikawa, H.; Joshikawa, H.; Joshikawa, H.; Joshikawa, H.; Danchin, A.; Toshikawa, H.; Joshikawa, H.; Joshikawa, H.; Joshikawa, H.; Danchin, A.; Toshikawa, H.; Joshikawa,
 A; Accession: H69719 A; Status: nucleic acid sequence not shown; translation not shown A; Status: nucleic acid sequence not shown; translation not shown A; Molecule type: DNA A; Molecule type: DNA A; Residues: 1-56 < KUN> A; Experimental source: GB: 299115; GB: AL009126; NID: 92634478; PIDN: CAB14066.1; PID: e11835 A; Experimental source: strain 168 C; Genetics: <la1> A; Genetics: <la1> C; Genetics: <ku1></ku1></la1></la1>
 Gene: sunA ;Superfamily: unassigned lanthionine-containing peptides; ;Keywords: antibiotic; lanthionine ;1-19/Domain: propeptide #status predicted <pro> ;20-56/Product: sublancin 168 #status predicted <mat> ;26-55/Disulfide bonds: #status experimental ;33-48/Disulfide bonds: #status predicted ;35/Modified site: dehydroalanine (Ser) #status experimenta ;38-41/Cross-link: (2S,3S,6R)-3-methyl-lanthionine (Thr-Cys</mat></pro>
 Query Match Best Local Similarity 100.0%; Pred. No. 4.4e-22;

ö

Gaps

; 0

Indels

385;

Length

```
C; Species: Homo sapiens (man)
C; Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 24-Nov-1999
C; Date: 04-Oct-1991 #sequence_revision 04-Oct-1991 #text_change 24-Nov-1999
C; Accession: A39248; A30328; A40109; A60946; A34942; A27653; A40108; A40494; A32224;
R; Lubahn, D.B.; Brown, T.R.; Simental, J.A.; Higgs, H.N.; Migeon, C.J.; Wilson, E.M.; Proc. Natl. Acad. Sci. U.S.A. 86, 9534-9538, 1989
A; Title: Sequence of the intron/exon junctions of the coding region of the human andr A; Reference number: A39248; MUID: 90083302
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           A; Molecule type: DNA
A; Residues: 1-919 <LUB>
A; Residues: 1-919 <LUB>
A; Cross-references: GB: M27423; GB: M27430; NID: g178904; PIDN: AAA51886.1; PID: g178906
B; Faber, P.W.; Kuiper, G.G.J.M.; van Rooij, H.C.J.; van der Korput, J.A.G.M.; Brinkma Mol. Cell. Endocrinol. 61, 257-262, 1989
A; Title: The N-terminal domain of the human androgen receptor is encoded by one, larg A; Reference number: A30328; MUID: 89137730
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               of human and rat complementary DNA encoding androgen recep; MUID:88178111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a cDNA encoding the human androgen rece
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                complementary DNA and localization to
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 559-624 <LU2>
A; Cross-references: GB: M20132
B; Kuiper, G.G.J.M.; Faber, P.W.; van Rooij, H.C.J.; van der Korput, J
J. Mol. Endocrinol. 2, R1-R4, 1989
A; Title: Structural organization of the human androgen receptor gene.
A; Reference number: A60946; MUID: 89322749
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Molecule type: DNA
A; Molecule type: DNA
A; Residues: 536-540;587-591;626-631;723-726;770-774;814-818;867-870 <
A; Residues: 536-540;587-591;626-631;723-726;770-774;814-818;867-870 <
A; Lubahn, D.B.; Joseph, D.R.; Sar, M.; Tan, J.; Higgs, H.N.; Larson,
Mol. Endocrinol. 2, 1265-1275, 1988
A; Title: The human androgen receptor: complementary deoxyribonucleic
A; Reference number: A34942; MUID: 89112208
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIDN:AAA51729.1; P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Molecule type: mRNA
A; Molecule type: mRNA
A; Residues: 468-564, 'K', 566-919 <TRA>
A; Cross-references: GB:M20260; NID:g178891; PIDN:AAA51774.1;
A; Note: the authors translated the codon AAG for residue 565
R; Chang, C.; Kokontis, J.; Liao, S.
Science 240, 324-326, 1988
A; Title: Molecular cloning of human and rat complementary DNA
A; Reference number: A40108; MUID:88178111
                                                                                                                                         A; Molecule type: DNA
A; Residues: 1-77,79-165,'A',167-389,'L',391-464,473-538
A; Cross-references: GB: M20260
R; Lubahn, D.B.; Joseph, D.R.; Sullivan, P.M.; Willard, H
Science 240, 327-330, 1988
A; Title: Cloning of human androgen receptor complementar
A; Reference number: A40109; MUID: 88178112
A; Accession: A40109
                                                                                                                CASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGG
                                     77;
16;
    Score 151; DB 2;
Pred. No. 5.7e-07
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A; Cross-references: GB:M20132; NID:g178627; PIDN:PR; Trapman, J.; Klaassen, P.; Kuiper, G.G.J.M.; var.Biochem. Biophys. Res. Commun. 153, 241-248, 1988 A; Title: Cloning, structure and expression of a cEA; Reference number: A27653; MUID:88240407 A; Accession: A27653
                                                      Mismatches
                                                      0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A; Molecule type: mRNA
A; Residues: 557-628 <CHA>
A; Cross-references: GB:M18624
R; Chang, C.; Kokontis, J.; Liao, E
Proc. Natl. Acad. Sci. U.S.A. 85,
27.78;
63.68;
                                                        Conservative
                                                                                                                                                                                                                                                                                                       - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Accession: A34942
Molecule type: mRNA
Residues: 1-919 <LU3>
                                Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Reference number: A; Accession: A60946
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Accession: A40108
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: A30328
                                                                                                                                                                                                                                                                                                        androgen receptor
                                                        28;
       Query Match
Best Local S
Matches 28
                                                                                                                33
                                                                                                                                                                    141
                                                                                                                                                                                                                                                     4
                                                                                                                                                                Db
                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                  Kidaris, Jan.

Kidecession: S18750

A. Molecule type: DNA

A. Residues: 1-336 < Lab.

Kiderences: EMBL:X59995

Kidericas references: EMBL:X59995

R. Davis, J. M.; Clarke, H.R.G.; Bradshaw Jr., H.D.; Gordon, M.P.

Plant Mol. Biol. 17, 631-639, 1991

A. Title: Populus chitinase genes: structure, organization, and similarity of translated A. Reference number: S17755, MUID:92003678

A. Reference number: S17755, MUID:92003678

A. Reference number: S17755

A. Status: nucleic acid sequence not shown

A. Residues: 1-103 < LoAv2>

A. Cross-references: EMBL:X59995

C. Genetics:

C. Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl.

C. Superfamily: lectin-related plant chitinase; hydrolase; polysaccharide degradation

F. 1-21/Domain: signal sequence #status predicted < Mary

F. 22-336/product: chitinase #status predicted < Mary

F. 22-336/promain: spacer

F. 83-336/Domain: catalytic #status, predicted < CAT>

F. 88-326/Domain: catalytic #status, predicted < CAT>

F. 88-326/Domain: plant chitinase homology < PCH>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ď
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     translated
                                                                                                                                                                                                                                                                               cottonwood)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5
      ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CESP: E02A10.2
                                                                                                                                                                                                                                                                            m poplar x
0-Sep-1999
          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -- CGGGAVACONYRQFCRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -CDGG
                                                                                                                                                                                                                                                     nwood
                                                           56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                           MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                poplar x cotton
(western balsam
#text_change 10
                                                                                       LFLSLSVRGSAEQCGQQAGDALCPG-GLCCSSYGWCGTTADYCGDG---CQSQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length 336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GSPDB:GN00023;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     hypothetical protein E02AlO.2 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change
C; Accession: T2O410
R; Thomas, K.
Submitted to the EMBL Data Library, October 1996
A; Reference number: Z19271
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 1; L
2.4e-08;
Thes 33;
    ö
                                                                                                                                                                                                                                             intinase (EC 3.2.1.14) precursor - western balsam
Species: Populus trichocarpa x Populus deltoides
Date: 10-Sep-1999 #sequence_revision 10-Sep-1999
Accession: $18750; $17755
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Reference number: 219271
A; Accession: T20410
A; Status: preliminary; translated from GB/EMBL/DDB
A; Molecule type: DNA
A; Residues: 1-385 <WIL>
A; Cross-references: EMBL: 281053; PIDN: CAB02877.1;
A; Experimental source: clone E02A10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                220/3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG~
      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 d. No. 2.4 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             85/1; 122/1; 133/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred
    ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30.3%;
Similarity 42.4%;
42; Conservative 6
    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             32/1; 72/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CESP: E02A10.2
   56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Gene: CESP:E
A; Map position
                                                                                                                                                                                                                                                                                                                                                       J.M.
                                                                                                                                                                                                                  S18750
chitinase (EC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                               C; Date: 10-Se
C; Accession:
R; Davis, J.M.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C; Genetics
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Introns:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                11
                                                         -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              4
   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
T20410
                                                                                                          a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ð
                                                         ð
```

the

Ris-Sta

.A.G.M.;

כי

ы .

F.S.; Wilson,

H.F.; French,

<FAB>

S

acid cloning

Œ,

R.E.; French,

<KUI>

PID:g178628; GB:J03180 ;, J.A.G.M.; Faber, P.W.;

PID:g178892 as Glu

1988

7211-7215,

Thu

```
Best Local Similarity
Matches 30; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
nes 30; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          85 GSGSG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match
Best Local S
Matches 30
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ŏ
                                                                                                                                                                                                                                                                                S.; Mauvai
                                                                                                                                                                                                                                                                                                                               and
and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   H
                                                                                                                                                               human androgen receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Jesse,
                                                                                                                                                                                                                                                                                                                            the DNA-binding domain of the
 human
                                                                                                                                                                                                                       A; Molecule type: mRNA
A; Residues: 1-77,79-211,'R',213-471,473-919 <TIL>
A; Residues: 1-77,79-211,'R',213-471,473-919 <TIL>
A; Cross-references: GB:M21748; GB:J04150; NID:g178871; PIDN:AAA51771.1; PID:g178872
A; Cross-references: GB:M21748; GB:J04150; NID:g178871; PIDN:AAA51771.1; PID:g178872
B; MOWSZOWicz, I.; Lee, H.J.; Chen, H.T.; Mestayer, C.; Portois, M.C.; Cabrol, S.; Mi Mol. Endocrinol. 7, 861-869, 1993
A; Title: A point mutation in the second zinc finger of the DNA-binding domain of the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Arabidopsis thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ng protein homology on; zinc finger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Α.;
seguences of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           14-May-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Duesterhoeft,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       54
                                                                                  646-919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         6
acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A; Introns: 538/2; 589/1; 628/1; 724/1; 772/2; 816/1; 868/3
C; Superfamily: unassigned erbA-related proteins; erbA transformis
C; Keywords: DNA binding; steroid binding; transcription regulation
F; 557-815/Domain: erbA transforming protein homology <ERBA>
F; 559-579/Region: zinc finger
F; 595-619/Region: zinc finger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          --AVACONYROF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            KLENPLDYGSAWAAAAA --- OCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 221;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             glycine-rich cell wall structural protein homolog F23E13.120 - PC; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change C; Accession: T04592
R; Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Due submitted to the Protein Sequence Database, March 1998
A; Reference number: 215378
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                       A; Molecule type: mRNA
A; Residues: 1-74,79-89,'H',90-472,'GGG',473-474,'E',476-644,
A; Cross-references: GB:M23263
A; Cross-references: GB:M23263
R; Tilley, W.D.; Marcelli, M.; Wilson, J.D.; McPhaul, M.J.
Proc. Natl. Acad. Sci. U.S.A. 86, 327-331, 1989
A; Title: Characterization and expression of a cDNA encoding A; Reference number: A32224; MUID:89098909
  amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Experimental source: cultivar Columbia; BAC clone F23E13
C; Genetics:
A; Map position: 4
A; Note: F23E13.120
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  34 ASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGGGGG
 and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 142.5; DB 2, Pred. No. 7.3e-06; 3; Mismatches 22,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GTIGCGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 141; DB 2;
Pred. No. 3.1e-06,
1; Mismatches 14
                                                                                                                                                                                                                                                                                                                                                                                                                translation
Structural analysis of complementary DNA nce number: A40494; MUID:89017168
                                                                                                                                                                                                                                                                                                                                                                                   A; Status: not compared with conceptual translative A; Molecule type: DNA
A; Residues: 557-614, 'H', 616-624 <MOW>
C; Genetics: PIDN: AAB28340.1; PID:g425580
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Cross-references: GDB:120556; OMIM:313700
A; Map position: Xq11-Xq12
A; Introns: 538/2; 589/1; 628/1; 724/1; 772/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            A40715; MUID:94019395
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               э;
Э
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ELENOKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25.9%;
illarity 65.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL: AL022141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26.1%;
42.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 42.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     <BEV>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 28; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A; Reference number:
A; Accession: T04592
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     A;Residues: 1-221 <FA;Cross-references:
                                                                                                                                                                                                                                                                                                                                                                      A; Reference number:
A; Accession: A40715
                     A; Reference number: A; Accession: A40494
                                                                                                                                                                                                       A; Accession: A32224
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  1-221
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Molecule type:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GDB: AR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Mar.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 A; Gene:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ğ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ò
```

```
gene
RESULT 6
T20268
hypothetical protein C56A3.1 - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C; Date: 15-Oct-1999

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          enhanced genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      a rice sucrose phosphate synthase (SPS)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          H.; Fujimura,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          glycine-rich protein Tfm5 - tomato
c; Species: Lycopersicon esculentum (tomato)
C; Species: Lycopersicon esculentum (tomato)
C; Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000
C; Accession: T07381
R; Santino, C.G.; Stanford, G.L.; Conner, T.W.
Plant Mol. Biol. 33, 405-416, 1997
A; Title: Developmental and transgenic analysis of two tomato fruit enhanced g
A; Reference number: 216000; MUID:97201476
A; Reference number: 216000; MUID:97201476
A; Accession: T07381
A; Accession: T07381
A; Molecule type: DNA
A; Molecule type
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sucrose-phosphate synthase (EC 2.4.1.14) 1 - rice
C; Species: Oryza sativa (rice)
C; Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C; Accession: T04103
R; Sakamoto, M.; Satozawa, T.; Kishimoto, N.; Higo, K.; Shimada, H.; Fujimura
Plant Sci. 112, 207-217, 1995
A; Title: Structure and RFLP mapping of a rice sucrose phosphate synthase (SF
A; Reference number: Z15212
A; Accession: T04103
A; Status: preliminary; translated from GB/EMBL/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -666666666666666666 121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 138; DB 2; Pred. No. 5.7e-06; l; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 139.5; DB Pred, No. 7e-06; ); Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Superfamily: hydroxyproline-rich glycoprotein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -GGGSGTGGGGGS-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              999939--99993999--99999--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25.3%;
larity 50.8%;
Conservative 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25.6%;
llarity 50.8%;
Conservative
```

Ħ

-09-894-030-3

Ŋ

<MAT>

ä

```
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter Nature 402, 761-768, 1999
A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana. A; Reference number: A84420; MUID: 20083487
A; Reference number: A84420; MUID: 20083487
A; Accession: B84777
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-255 <STO>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                carboxyl terminal homolo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              encoding glycine-rich
                                                                                                           #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 GSPDB:GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              hypothetical protein At2g36120 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              glycine-rich protein GRP22 - rape
C; Species: Brassica napus (rape)
C; Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C; Accession: S31415
                                                                                                                                                                                                                                                                                                                                                                                                                                            A; Status: preliminary
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-291 <BER>
A; Cross-references: EMBL: Z15045; NID: g17820; PIDN: CAA78762.1; PID: g17821
C; Superfamily: Phaseolus glycine-rich cell wall protein 1.8
                                                                                                                                                                                                                                                                                                                                                                                                                     --FCRGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                                                        .;
8
                                                                                                                                                                                                                                                                                                                        Length 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 255;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GB:AE002093; NID:g4678224; PIDN:AAD26969.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              a gene family
     A; Gene: grp-1
C; Superfamily: glycine-rich cell wall structural protein
C; Keywords: cell wall; duplication; structural protein
C; Keywords: cell wall; duplication; structural protein
F; 1-23/Domain: signal sequence #status predicted <SIG>
F; 24-165/Product: glycine-rich cell wall structural protein 1
F; 30-55/Region: repeat R1
F; 56-62/Region: repeat R2
F; 62-92/Region: repeat R1
F; 93-99/Region: repeat R2
F; 100-131/Region: repeat R1
F; 139-160/Region: repeat R1
                                                                                                          7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A;Gene: At2g36120
A;Map position: 2
C;Superfamily: collagen alpha 1(V) chain; fibrillar collagen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Bellemare,
                                                                                                                                                                                                                                                                                                                                                                     22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 130; DB 2; I
Pred. No. 3.8e-05;
); Mismatches 19;
                                                                                                                                                                                                                                                                                                                       Score 130; DB 1;
Pred. No. 2.7e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 C; Accession: S31415
R; Bergeron, D.; Boivin, R.; Baszczynski, C.L.; Belsubmitted to the EMBL Data Library, August 1992
A; Description: Characterization and expression of A; Reference number: S31415
A; Accession: S31415
                                                                                                                                                                                                                                                                                                                                                                                                                   18 GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQ--
                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ;
0
                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                       23.98;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23.9%;
50.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Conservative
                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 30; Conserv
                                                                                                                                                                                                                                                                                                                                             Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          A;Residues: 1-255 < A;Cross-references:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | | | | | | GSGEGGAH-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 73 GGGG 76
                                                                                                                                                                                                                                                                                                                                                                 31;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              156 GGSG
                                                                                                                                                                                                                                                                                                                       Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            182
                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Qy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
                                                                                                                                                                                                                                                                                                                                                                                                                   δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PIDN:CAA46951.1; PID:g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Boncinelli, E.; Krumlau
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ve transcripts with di-
A; Residues: 1-1084 <SAK>
A; Cross-references: EMBL:D45890; PIDN:BAA08304.1
A; Experimental source: subsp. Japonica
C; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Genetics:
A; Map position: 1
A; Introns: 120/3; 200/2; 221/3; 452/3; 492/3; 536/3; 595/3; 613/3; 634/3; 946/3; 989/2
C; Superfamily: sucrose-phosphate synthase; sucrose/sucrose-phosphate synthase homology
C; Keywords: glycosyltransferase; hexosyltransferase
E; 196-680/Domain: sucrose/sucrose-phosphate synthase homology <SSPS>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16-Jul-1999
                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (clone lambda-313)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PID:920247
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      N.Alternate names: homeotic protein Hox 2.7
C; Species: Mus musculus (house mouse)
C; Date: 22-Nov-1993 #sequence_revision 21-Jul-1995 #text_change 20-Au.
C; Accession: $20963; D42694
R; Sham, M.H.; Hunt, P.; Nonchev, S.; Papalopulu, N.; Graham, A.; Bonc EMBO J. 11, 1825-1836, 1992
A; Title: Analysis of the murine Hox-2.7 gene: conserved alternative t: A; Reference number: $20963; MUID:92258392
A; Reference number: $20963; MUID:92258392
A; Residues: 1-433 cSHA>
A; Residues: 1-433 cSHA>
A; Residues: 1-433 cSHA>
A; Residues: 1.433 cSHA>
A; Residues: 213 cSHA>
A; Residues: 213 cSHA>
A; Reference number: A42694; MUID:92212934
A; Reference number: A42694; MUID:92212934
A; Residues: 213 cSHA cNAZ>
A; Residues: 213 cSHA cNAZ>
A; Residues: 213 cSHA comeobox homeobox homeology
C; Reywords: DNA binding; homeobox; nucleus; transcription regulation F; 192-248/Domain: homeobox homeology <HOX>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Residues: 213-238 <NAZ>
Note: sequence extracted from NCBI backbone (NCBIN:92310, NCBIP:923 Superfamily: homeotic protein Hox B3; homeobox homology Keywords: DNA binding; homeobox; nucleus; transcription regulation 192-248/Domain: homeobox homology <HOX>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            15;
                                                                                                                                                                                                                                                                         1084;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    433;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               glycine-rich cell wall structural protein 1 precursor (clone lam C; Species: Oryza sativa (rice)
C; Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change
C; Accession: S13385
R; Lei, M.; Wu, R.
Plant Mol. Biol. 16, 187-198, 1991
A; Title: A novel glycine-rich cell wall protein gene in rice.
A; Reference number: S13385; MUID: 91370862
A; Accession: S13385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -GGGGGGGGGMSK
                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PIDN:CAA37665.1;
                                                                                                                                                                                                                                                                                                                                                                    11
                                                                                                                                                                                                                                                                                                                                                                                                                  46
                                                                                                                                                                                                                                                                                                                                                                                            -666666666666666666V
                                                                                                                                                                                                                                                                                                                                                                SGGTIGCGGGAVACONYROFCRGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 132.5; DB 2;
Pred. No. 3.4e-05;
2; Mismatches 6;
                                                                                                                                                                                                                                                                   Score 134.5; DB 2;
Pred. No. 4.8e-05;
L; Mismatches 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              31 LQCASGGTI-GCGGGAVACQNYRQFCRGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A;Molecule type: DNA
A;Residues: 1-165 <LEI>
A;Cross-references: EMBL:X53596; NID:g20246;
C;Genetics:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2;
                                                                                                                                                                                                                                                                                                                   7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 24.3%;
ilarity 56.6%;
Conservative
                                                                                                                                                                                                                                                                   24.7%;
llarity 60.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LKNSSPGTAEGCGGGG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Query Match
Best Local Similarity
Watches 30; Conser
                                                                                                                                                                                                                                                                                          Similarity
26; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           S20963
homeotic protein Hox
                                                                                                                                                                                                                                                                                                                                                                                                                SGGAAGGGG-
                                                                                                                                                                                                                                                                     Query Match
Best Local 3
                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                              17
                                                                                                                                                                                                                                                                                                                                                                 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Φ
                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              δ
```

ä

230

d

g

δ

```
C; Species: Homo saplens (man)
C; C; Date: 10-Sep-1991 sequence_revision 30-Sep-1991 ftext_change 22-Jun-1999
C; Accession: S7543; S1547; D37042
C; Accession: S7543
C; Mucleic Acids Res. 17, 1085-10402, 1989
A; Title: The human HOX gene family
A; Recteance number: S07541; MUD: 90098876
A; Accession: S07543
A; Accession: S07543
A; Residues: 1-431 - Acx
A; Residues: 1-431 - Acx
A; Cross references: EMBL: X16667; NID: 932379; FIDN: CAA34657.1; PID: 932380
A; Title: Organization of human class I homeobox genes.
A; Reference number: S15036; MUD: 90215256
A; Accession: S15547
A; Molecule type: DMX
A; Residues: 188-253 - CBON
A; Rectence number: A37042; MUD: 932377; PIDN: CAA34297.1; PID: 9330068
C; Genetics:
A; Residues: 180-253 - CBON
A; Residues: 180
                                                                                                                                                                                                                                      EF hand; heterodimer;
A; Cross-references: GDB:119752; OMIM:114170
A; Map position: 19pter-19qter
A; Introns: 70/2; 81/3; 111/3; 131/1; 152/3; 175/3; 202/1; 241/1; 260/3
C; Complex: heterodimer of L (large) and S (small) chains
C; Complex: heterodimer of L (large) and S (small) chains
C; Function:
A; Description: catalyzes the hydolysis of peptides
C; Function:
A; Note: cleaves preferentially after tyrosine, methionine, or arginine residues and
C; Superfamily: calpain small chain; calmodulin repeat homology
C; Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer;
C; Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer;
F; 1-56/Domain: calmodulin repeat homology <EF1>
F; 139-171/Domain: calmodulin repeat homology <EF2>
F; 204-236/Domain: calmodulin repeat homology <EF5>
F; 237-268/Domain: calmodulin repeat homology <EF5>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N; Alternate names: homeotic protein Hox 2.7; homeotic protein Hox 2G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             -GGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                   ; DB 1; Lot
4.9e-05;
ohes 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 129; DB 1; Pred. No. 7.3e-05; h; Mismatches 7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pred. No. 4.96
0; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 129;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                          th 23.7%; Similarity 50.0%; 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23.7%;
larity 58.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   - human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SPGTAEGCGGGG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     homeotic protein Hox B3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity 27; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      146
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               human calcium-dep
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Town, C.D.; Fujii, C.Y.;
I.; Umayam, L.; Tallon, L.; Fraser, C.M.; Venter,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      neutral
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               bidopsis thaliana
                                                                                            3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    calgain (EC 3.4.22.17) small chain - human
N;Alternate names: calcium-activated neutral proteinase (CANP)
C;Species: Homo sapiens (man)
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 16-Jul-1999
C;Accession: A26107; A23650
C;Accession: A26107; A23650
C;Accession: A26107; A3648; MulD:87066759
A;Title: Gene organization of the small subunit of human calcium-activated neut
A;Title: Gene organization of the Small subunit of human calcium-activated neut
A;Reference number: A93648; MulD:87066759
A;Accession: A26107
A;Molecule type: DNA
A;Residues: 1-268 <MIY>
A;Cross-references: GB:M31502
R;Ohno, S.; Emori, Y.; Suzuki, K.
Nucleic Acids Res. 14, 5559, 1986
A;Title: Nucleotide sequence of a cDNA coding for the small subunit of human calcumpers of a cDNA coding for the small subunit of human calcumpers.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GSPDB:GN00139
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        02-Feb-2001
                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PID:935328
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18 GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNY----RQFCRGGGGGGGGGGGGGGGGGGGGGGGGG
                                                                                       16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       hypothetical protein At2g37830 [imported] - Arabidopsis thaliana ("Species: Arabidopsis thaliana (mouse-ear cress)
C; Species: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change (C; Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change (C; Accession: F84797
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; To M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Nature 402, 761-768, 1999
A; Title: Sequence and analysis of chromosome 2 of the plant ArabiA; Reference number: A84420; MUID: 20083487
                                     Length 291
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GB:AE002093; NID:g4895201; PIDN:AAD32788.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Cross-references: EMBL: X04106; NID: 935327; PIDN: CAA27726.1; C; Genetics: A; Gene: GDB: CAPN4
                                                                                            Indels
                                 ; DB 1; L
4.3e-05;
ches 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23.7%; Score 129; DB 2; 44.4%; Pred. No. 2.3e-05;
                                     Score 130; DB Pred. No. 4.3e-2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ed. No. 2.36
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --GDGGGGGFGGGG--
                                                                                       2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ij
                                 23.9%;
50.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Molecule type: DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A; Accession: A23650
A; Molecule type: mRNA
A; Residues: 1-268 < OHN>
                                                              Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              preliminary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   64 DGANGGGFG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Cross-references: G; Genetics:
A; Gene: At2g37830
A; Map position: 2
                                                                                                                                                                                                        221 GSGAGGAH--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: F84797
A; Status: prelimina
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    102 GGG 104
                                                                                                                                                                                                                                                                                                                     269 GG 270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               28;
                                                                                    31;
                                                                                                                                                                                                                                                               74 GG 75
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
                                   Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                       Matches
```

q

ð

g

à

¥.;

ä

Ľτι

Search completed: October 24, 2002, 09:29:44 Job time : 20.9781 secs .

```
Ltd
5.1.3
Compugen
  GenCore version (c) 1993 - 2002
         Copyright
```

mode1 S protein search, using OM protein

Seconds 2002, 09:28:26 October 24, Run on:

updates/sec ; Search time 56.2044 (without alignments) 531.492 Million cell

US-09-894-030-3 score: Perfect

.... DDFDLDVVKVSKQDSKITPQ 100 1 MEKLFKEVKLEELENQKGSG.... Sednence:

BLOSUM62 Gapop 10.0 , Gapext Scoring table:

298721915 residues 1105779 segs, Searched:

0.5

1105779 satisfying chosen parameters: hits of number Total

Minimum

seq length: 0 seq length: 2000000000 80 68 68 Maximum

Post-processing: Minimum Match 100% Maximum Match 100% Listing first 45 s

summaries

Database

Pending\_Patents\_AA\_New:\*

'cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*

'cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\* 1264597

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		;	A	Ą	A	Ą	Ą	A	p]	A	p]	dd	A	A	pj	Ø	Ą	Ą	Ą	Ø	A	Ø	Ą	Ø	A	V	A	Ø
		1	145	742	684	045	488	49	6, Ap	5459,	6, Ap	12, A	8475,	29	0, Ap	8976,	007	631	631	573	70	035	531	6550,	61	545	413	764
	+ 100	<b>ا</b> ا	e 3	e 5	е 9	e 8	7	7	m	æ	m	~	7	~	m	æ	N	~	N	C	7	-	N	~	7	e 5	9	ce 4
		4 1	edneuc	ednen	equenc	equence	ednence	ednence	quence	ednence	ednence	ednence	equence	edneuc	equence	ednence	equence	ednence	equence	ednence	edneuce	edneuce	ednence	ednence	quenc	equenc	quenc	nen
	Ū	חמטמ		Sec	Sec		Sec	Sec	U	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec	Sec			Sec	Ü	Sec	Sec
		1 1 1 1			91					63			75	07										20				
			31	574	968	804	748	749	3	85		21	784	29	3	89	00	$^{\circ}$	63	7	87	$\sim$	53	65	9	54	H	9
ES		1	5	53	S	53	53	53	39	53	823	16	53	537	18	537	995	134	136	993	997	866	135	137	138	139	537	-666-
SUMMARIES		1	79	79	791-	79	79	79	0	79	20	25	79	79	94	79	18	20	20	18	18	18	20		20	20	79	219
SUM		 	-0	9	-0	0-	9	-60-	Ţ	-0	-1	7	9	9	9	-0	-1	•	7	7	7	ı	-	7	7	7	-60-	-1
	Ļ	7 :	ns	ΩS	ns	US	ns	nS	$^{PC}$	ns	OS	nS	ns	us	us	nS	US	nS	ns	$\Omega$ S	nS	US	SD	ns	ns	US	ns	ns
	g		'n	വ	Ŋ	Ŋ	S	Ŋ	Ţ	Ŋ	9	ဖ	വ	Ŋ	വ	ഹ	9	9	9	9	9	9	9	9	9	9	S	9
	Tonath	7 6 7 7 7	303	ß	3	3	Н	918	$\vdash$	Ч	Η,	8	Ч	3	4	$\sim$	0	0	0	0	0	0	0	0	0	O	4	0
	ery	1 1	•	•				•			•				•			•	•	•	•	•	•	•	•	•	٠	•
ф	One	ING.	ĕ	ĕ		~	Ñ	Ä	5	Š	Ñ	Ñ	~	₹	Ň	Ä	Ċ	Ò	Ö	Ö	Ö	C)	Ä	7	<b>(</b> 2	Ò	7	7
	Ç	3000	165	9	4	42.	42.	•	42.	42.	42.	41.	4	3	•	$\sim$	34.	34.	34.	34.	34.	34.	34.	•	34.	34.	34.	34.
			1	7	m	4	ស	9	7	8	თ	10	11	12	13	14	15	16	17	1.8	19	20	21	22	23	24	25	26

Sequence 9, Appli Sequence 531, App Sequence 531, App	equence 531, Pequence 28229, equence 27910, equence 25833, equence 309, Pequence 3009, Pequence 309, Pequence 3009, Pequence 300
US-10-217-700-9 US-10-123-155-531 US-10-137-871-531 US-10-141-761-531 US-10-140-864-531 US-10-140-923-531 US-10-141-759-531 US-10-140-472-531 US-10-142-885-531 US-10-142-885-531	-10-158-790-5 -10-155-881-2 -10-155-881-2 -10-155-881-2 -10-133-155-3
αουουουουου	००००००
1084 1150 1150 1150 1150 1150 1150 1150	
444444444444444444444444444444444444444	· · · · · · · · · · · · · · · · · · ·
134.5 134.5 134.5 134.5 134.5 134.5 134.5 134.5	
22 22 33 33 34 34 35 36 37	44444 44444 5432100

#### ALIGNMENTS

```
APPLICANT: Bionomix, Inc.
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
APPLICANT: Deseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT APPLICATION NUMBER: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 57429
LENGTH: 336
                                                                                                                                                  Σ
                                                                                                                                                  AND FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                            5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             63
                                                                                                                                                PROTEIN FAMILIES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --- CGGGAVACONYRQFCRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16;
                                                                                                                                                                                                                                                                                                                                                                                                                    Length 303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
RESULT 1
US-09-791-537-31453
; Sequence 31453, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF
; TITLE OF INVENTION: METHODS OF USE THEREOF
; TITLE OF INVENTION: METHODS OF USE THEREOF
; TITLE PREFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 31453
; LENGTH: 303
                                                                                                                                                  OF
                                                                                                                                                                                                                                                                                                                                                                                                                   Score 165; DB 5; 1
Pred. No. 2.8e-07;
8; Mismatches 33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               86
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGGGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4 LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 2
US-09-791-537-57429
; Sequence 57429, Application US/09791537
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                       ; ORGANISM: Populus balsamifera subsp
US-09-791-537-31453
                                                                                                                                                                                                                                                                                                                                                                                                                                                             .
ώ
                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity 42.4%;
Matches 42; Conservative
                                                                                                                                                                                                                                                                                                                                      TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ά
```

g

qq

Qγ

```
SSULT 5
S-09-791-537-74882
Sequence 74882, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PROTEIN FAMILIES AND FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4;
                                                  254
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  438
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   441
                   54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    54
                                                  KLENPLDYGSAWAAAAA - - - OCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 KLENPLDYGSAWAAAA---OCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 382 KLENPLDYGSAWAAAA---QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 918;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ----GTIGCGGG---AVACONYRQF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -GTIGCGGG---AVACONYRQF-
                    -- AVACONYROF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    OF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 142.5; DB 5;
Pred. No. 8.1e-05;
3; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142.5; DB 5;
Pred. No. 8.1e-05;
3; Mismatches 22;
                   -GTIGCGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRUCTURES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCT
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
                                                                                                                                                                                                                                                                                                            CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TITLE OF INVENTION.
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
NUMBER OF SEQ ID version 3.0
                                                                                                              281
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             468
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 74905, Application US/09791537 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26
                                                                                                QLYGPCGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --- CRGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  -CRGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                -CRGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12 ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                   ELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                           SOFTWARE: Patentin version 3.0 SEQ ID NO 74882 LENGTH: 917
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26.18;
42.58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        26.1%;
42.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SOFTWARE: Patentin version
SEQ ID NO 74905
LENGTH: 918
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                             sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Local Similarity
Matches 37; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                         ; ORGANISM: HOMO
US-09-791-537-74882
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; ORGANISM: Homo US-09-791-537-74905
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
Matches 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                  198
                                                                                55
                                                                                                             255
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  382
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 55
                                                                                                                                                            RESULT
                                                                                                                                                                          us.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  d
                                                  ದ್ದ
                                                                                                             Dp
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ω
                    ρy
                                                                                δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δλ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 4
US-09-791-537-80458

Sequence 80458, Application US/09791537

GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
TITLE OF INVENTION: METHODS OF USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                               MEMBE
                                                                                                                                                                                                                                                                                                                                                   APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                               5;
                                                                                               Gaps
                                                                                                                           --- CGGGAVACQNYRQFCRGG 58
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -CGGGAVACQNYRQFCRGGGGGGGGGGGGGG 70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -CDGG
                                                                                               16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  734
                                                                                                                                            Length 336;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                             33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 147; DB 5; 1
Pred. No. 1.3e-05;
1; Mismatches 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 142.5; DB 5;
Pred. No. 6.6e-05;
3; Mismatches 22;
                                                                                                                                                                                        97
                                                             Score 165; DB 5
Pred. No. 3e-07;
3; Mismatches
                                                                                                                                                                                     GGGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKI
                                                                                                                           LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 80458
LENGTH: 734
                                                                                                                                                                                                                                                         RESULT 3
US-09-791-537-96846
; Sequence 96846, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
applicant: Debe, Derek
                   dsqns
                                                                                           8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QKGSGLGKAQCAALWLQCASGGTIG
; TYPE: PRT
; ORGANISM: Populus balsamifera
US-09-791-537-57429
                                                            30.3%;
iilarity 42.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local Similarity 54.5%;
Matches 36; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               26.1%;
ilarity 42.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: Oryza sativa -09-791-537-96846
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     sapiens
                                                            Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     76
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; ORGANISM: HOMO : US-09-791-537-80458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQ ID NO 96846
LENGTH: 333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9999999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       71 GGGGGG
                                                                                       42;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE: PRT ORGANISM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   : PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches.
                                                                                                                                                                                                                     64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           16
                                                                                                                            マ
                                                                                                                                                                                        59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  78
                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query
Best I
```

a

Ω

ŏ

g

ΜË

Thu

g

```
APPLICANT: Kamatkar, Shubhangi
APPLICANT: Wonsey, Angela M.
APPLICANT: Glatt, Karen
APPLICANT: Zhao, Xumei
APPLICANT: Zhao, Xumei
APPLICANT: Anderson, Dustin
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION,
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                54
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               KLENPLDYGSAWAAAA---QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  --GTIGCGGG---AVACONYROF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5;
22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 142.5; DB (
Pred. No. 8.1e-05;
); Mismatches 22
                                                                                                                                                                                                                                                                                                                                                                             FILE REFERENCE: MRI-044

CURRENT APPLICATION NUMBER: US/10/205,823

CURRENT FILING DATE: 2002-07-25

PRIOR APPLICATION NUMBER: 60/307,982

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2001-12-12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    469
                                   469
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR FILING DATE: 2002-03-05
NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSEQ for Windows Version
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sequence 212, Application US/10259165 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     443 QLYGPCGGGGGGGGGGGGGGGGGG
                                                                                                                 Application US/10205823
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -CRGGGGGGGGGGGGGGGGGG
3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12 ELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                Gannavarapu, Manjula
Gorbatcheva, Bella
Hoersch, Sebastian
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Moughamer, Todd
Provart, Nicholas
Ricke, Darrell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Fumiyaki
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Wang, Xun
Chang, Hur-song
Briggs, Steven P.
Cooper, Bret
Glazebrook, Jane
Goff, Stephen A.
Katagiri, Fumiyak
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   th 26.1%;
Similarity 42.5%;
37; Conservative
                                                                                                                                                                                                                                  Hoers...
Kamatkar, Shub...
Angela M.
                                                                                                                                                                                  Wilson O
                                                                                                                                                  : Schlegel, Robert
: Monahan, John E.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Kreps, Joel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ; ORGANISM: Homo sapiens US-10-205-823-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tong
                                                                                                                                                                                    Endege,
                                                                                                  Sequence 36, Application GENERAL INFORMATION
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zhu,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 10
US-10-259-165-212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SOFTWANE.
SEQ ID NO 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local $
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICANT: APPLICANT:
                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                   APPLICANT:
                                                                                                                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          APPLICANT
                                                                                                                                                     APPLICANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
                                                                                      RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                     q
       δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FAMILIES AND FAMILY MEMBE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         4;
                                                                                                                                                       PREVENTION, AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KLENPLDYGSAWAAAA----QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAM
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 85459
LENGTH: 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -AVACQNYRQF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              --GTIGCGGG---AVACQNYRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          9
                                                                                   Sequence 36, Application PC/TUS0223913
GENERAL INFORMATION:
APPLICANT: Millennium Pharmaceuticals, Inc. et al.
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT,
TITLE OF INVENTION: THERAPY OF PROSTATE CANCER
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 142.5; DB 5;
Pred. No. 8.1e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142.5; DB 1;
Pred. No. 8.1e-05;
3; Mismatches 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -GTIGCGGG-
                                                                                                                                                                   FILE REFERENCE: MRI-044PC

FILE REFERENCE: MRI-044PC

CURRENT APPLICATION NUMBER: PCT/US02/23913

CURRENT FILING DATE: 2002-07-25

PRIOR APPLICATION NUMBER: 60/307, 982

PRIOR FILING DATE: 2001-07-25

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-08-22

PRIOR FILING DATE: 2001-09-25

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2001-12-12

PRIOR FILING DATE: 2002-03-05

PRIOR FILING DATE: 2002-03-05
                                                                                                                                                                                                                                                                                                                                                                                                                        4.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Sequence 85459, Application US/09791537
GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
         465
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ELENQKGSGLGKAQCAALWLQCASG----
                                                                                                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQ ID NOS: 455
SOFTWARE: FastSEQ for Windows Version
SEQ ID NO 36
LENGTH: 919
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          439 QLYGPCGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             э;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            12 ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           .;
m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26.1%;
illarity 42.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26.1%;
illarity 42.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            l Similarity
37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity 37; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ORGANISM: Homo :
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SULT 8
-09-791-537-85459
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ; ORGANISM: HOMM
PCT-US02-23913-36
                                                                            PCT-US02-23913-36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              TYPE: PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SOFTWARE:
SEQ ID NO 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             386
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            52
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                           RESULT
```

g

ò

ŏ

qq

g

ð

4,

AND

us-09-894-030-3.rapn

4

4 ;

```
Melton, Do
TITLE OF INVENTION: 7
NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Celeste,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDIUM TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 INFORMATION FOR SEQ
                                                                                                                                                                                                      Local Similarity
nes 33; Conser
                                                                                                                                                                                                                                                                                                                      74
                                                                                                                                                                                                                                                                                                                                                        GGGG 80
                                                                                                                                                                                                                                                                                                                                                                                                           RESULT 13
US-09-945-182-30
                                                                                                                                                                                                                                                                                                                         GGGG
                                                                                                                      PRT
                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                        16
                                                                                                                                                                                                                                                                                                                         71
                                                                                                                                                                                                                     Matches
                                                                                                                       TYPE:
                                                                                                                                                                                                          Best
                                                                                                                                                                                                                                                                                        đ
                                                                                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                                                                                                                          δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 12
US-09-791-537-22910
; Sequence 22910, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITI,E OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                APPLICANT: Debe, Derek
APPLICANT: Danzer, Joseph
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
TITLE OF INVENTION: METHODS OF USE THEREOF
FILE REFERENCE: 261/210
     SILENCING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                5;
    IONAL GENE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                         -CGGGAVACQNYRQFCRGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -0666
                                                                                                                                                                                                                                                                                                                                                                          16;
                                                                                                                                                                                                                                                                                                                                       17;
                                                                                                                                                                                                                                                                                                        280;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11 LFK----NSYAQQCGSQAGGALCSNR-LCCSKFGYCGSTDPYCGTG---CQSQ--
     POSTTRANSCRIPT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 318;
                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                        version 3.0.4 (C) 2001 Syngenta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 141; DB 5; Pred. No. 4.3e-05; S; Mismatches 22;
                                                                                                                                                                                                                                                                                                      26.0%; Score 141.5; DB 6; larity 46.2%; Pred. No. 3.4e-05; Conservative 3; Mismatches 15;
    THAT ARE MODULATED BY
                                                                                                                                                                                                                                                                                                                                                                          20 GLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG
TITLE OF INVENTION: GENES THAT ARE MODULATED FILE REFERENCE: 70030-NP CURRENT APPLICATION NUMBER: US/10/259,165 CURRENT FILING DATE: 2002-09-26 PRIOR APPLICATION NUMBER: US 60/370,620 PRIOR FILING DATE: 2002-04-04 PRIOR APPLICATION NUMBER: US 60/368,327 PRIOR FILING DATE: 2002-03-27 PRIOR FILING DATE: 2001-09-26 NUMBER OF SEQ ID NOS: 782 SOFTWARE: PatentList.pl version 3.0.4 (C) 20 SEQ ID NO 212 LENGTH: 280
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
SEQ ID NO 78475
LENGTH: 318
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-09-791-537-78475
; Sequence 78475, Application US/09791537
; GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                            25 GCQSNMCCSKWGYCGTGKDY-CGDG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              25.9%;
illarity 46.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     GGGGGGGGGGGGGGGMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Allium sativum
-78475
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
APPLICANT: Danzer, Josep
                                                                                                                                                                                                                                                        ; ORGANISM: Oryza sativa
US-10-259-165-212
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity 37; Conservation
                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                         74 GGGMS 78
                                                                                                                                                                                                                                                                                                                                                                                                                                                              | |:|
GSGVS
                                                                                                                                                                                                                                                                                                                                       30;
                                                                                                                                                                                                                                        PRT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ; ORGANISM:
US-09-791-537
                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9
                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                          a
a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   qa
                                                                                                                                                                                                                                                                                                                                                                                                                                           δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Б
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   δy
```

```
Gaps
                                                                                                                                                                                                                                                                              ----CGGGAVACQNYRQFCRGGGGGGGGGGGGGG
                                                                                                                                                                                                                           12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 240;
                                                                                                                                                                                          Length 335;
                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Rosen, Vic., Wolfman, Neil
Wolfman, Neil
Thomsen, Gerald H.
Melton, Douglas A.
Melton, Douglas A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DB 5;
                                                                                                                                                                                                                 5;
16;
                                                                                                                                                                                          Score 139; DB 5;
Pred. No. 6.8e-05;
3; Mismatches 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 135.5; DB Pred. No. 0.0001;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            INC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: 5202-D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617 498-8260
TELEFAX: 617 876-5851
'ORMATION FOR SEQ ID NO: 30:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/945,182
FILING DATE: 31-Aug-2001
CLASSIFICATION: <Unknown>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQ ID NO: 30:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ADDRESSEE: GENETICS INSTITUTE,
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: USA
FILE REFERENCE: 261/210
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: PatentIn version 3.0
SEQ ID NO 22910
LENGTH: 335
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/808,324
                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 30, Application US/09945182
GENERAL INFORMATION:
APPLICANT: Celeste, Anthony J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: <Unknown>ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                               QKGSGLGKAQCAALWLQCASGGTIG-
                                                                                                                                                                                                                           3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Cele.
Wozney, Jun.
                                                                                                                                         ; ORGANISM: Elaeagnus umbellata
US-09-791-537-22910
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   24.9%;
40.4%;
                                                                                                                                                                                          25.5%;
larity 51.6%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ZIP: 02140
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGIH: 240 amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TOPOLOGY: linear

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: 8
US-09-945-182-30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query
```

ъ,

Gaps

23;

Indels

Length 101;

2

---GGGGGGGG 47

```
6 KEVKLEELENOKGSGLGKAQCAALWLQCASGGTIGCGGGAVACONYROFCRGGGGGGG
SOFTWARE: Molecular Dynamics Sequence Listing Engine SEQ ID NO 20071
LENGTH: 101
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AP000161.1
FEATURE:
                                                                                                                                                                                                           Score 134.5; DB 6;
Pred. No. 5.9e-05;
3; Mismatches 15;
                                                                                                                                                                                                                                                                                                                         --99989999999---
                                                                                                                                                   ; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL US-10-182-995-20071
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Search completed: October 24, 2002, 09:35:11
Job time : 58.2044 secs
                                                                                                                                                                                                               Query Match
Best Local Similarity 44.6%;
Matches 33; Conservative
                                                                                                                                                                                                                                                                                                               66 GGGGGGGGGMSK 79
                                                                                                                                                                                                                                                                                                                                                                                                        61
                                                                                                                                                                                                                                                                                                                                                                                           GGGGGGGGGRRRK
                                                                                                                                                                                                                                                                                                                                                                                                           48
                                                                                                                                                                                                                                                                                                                                   QQ
                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                         δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN HEART
FILE REFERENCE: PB 0004 WO 1202-08-02
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 2002-08-02
PRIOR FILING DATE: 20 WHOMBER: US 60/207, 456
PRIOR FILING DATE: 20 May 2000 (25.05.00)
PRIOR FILING DATE: 20 May 2000 (25.05.00)
PRIOR PILING DATE: 03 checober 2000 (21.00)
PRIOR PILING DATE: 03 checober 2000 (21.00)
PRIOR PILING DATE: 21 September 2000 (21.09.00)
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR PILING DATE: 21 September 2000 (21.09.00)
PRIOR PILING DATE: 31 September 2000 (21.09.00)
PRIOR PILING DATE: 30 JUNE 2000 (31.00.00)
PRIOR PILING DATE: 31 September 2000 (31.09.00)
PRIOR PILING DATE: 31 September 2000 (31.09.00)
PRIOR FILING DATE: 30 JUNE 2000 (31.06.00)
PRIOR FILING DATE: 30 JUNE 2000 (31.06.00)
PRIOR FILING DATE: 30 JUNE 2000 (31.06.00)
                                                                                                                                                                                                                                                                                                                                                      STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5;
                4,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     106
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                   ---ELENQKGSGLG--KAQCAALWLQ-CASGGTIGCGGGAVACQN 50
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             14;
                23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 331;
                                                                              5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 135; DB 5;
Pred, No. 0.00015
                                                                                                                                                                                                                   RESULT 14
US-09-791-537-88976

US-09-791-537-88976

Sequence 88976, Application US/09791537

GENERAL INFORMATION:
APPLICANT: Bionomix, Inc.
APPLICANT: Debe, Derek
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTUR
TITLE OF INVENTION: METHODS OF USE THEREOF
TITLE OF INVENTION: METHODS OF USE THEREOF
CURRENT APPLICATION NUMBER: US/09/791,537
CURRENT FILING DATE: 2001-02-22
NUMBER OF SEQ ID NOS: 153055
SOFTWARE: Patentin version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ed. No. 0.00
Mismatches
                  Mismatches
                                                                                                                                                                    129
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          US-10-182-995-20071; Sequence 20071, Application US/10182995; GENERAL INFORMATION: APPLICANT: Molecular Dynamics, Inc.
                                                                                                                            11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        24.8%;
38.9%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               --SKITPQ 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GRGANISM: Mus musculus US-09-791-537-88976
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ::::|:
108 PGSPRLTEVSPE
                                                          EKLFKEVKLE-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQ ID NO 88976
LENGTH: 331
                     36;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                94 D---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT
                                                                                                                                   21
                                                                                              53
                                                          ~
                                                                                                                                                                      107
                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                      qq
                                                                                               g
                                                                                                                                    Q
                                                           ò
```

• 

.rsp

us-09-894-030-3

ч

```
5.1.3
Compugen Ltd.
 version
- 2002
GenCore
(c) 1993
         Copyright
```

using sw model - protein search, OM protein

Seconds ; Search time 11.6788 (without alignments) 331.536 Million cell u 09:24:01 2002, October 24, Run on:

updates/sec

Title: Perfect

US-09-894-030-3 545 1 MEKLFKEVKLEELENQKGSG.....DDFDLDVVKVSKQDSKITPQ 100 score: Sequence:

Scoring table:

0.5 BLOSUM62 Gapop 10.0 , Gapext

38719550 residues 105224 seqs, Searched:

105224 Total number of hits satisfying chosen parameters

2000000000 seq length: seq length: Minimum DB Maximum DB

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_40 Database

ing printed, have a Pred. No. is the number of results predicted by chance to score greater than or equal to the score of the result be and is derived by analysis of the total score distributio

## SUMMARIES

SUMMARIES	scription		ANDR_HUMAN P10275 homo sapien	mus musc	_MOUSE P09026	_HUMAN 06090	ORYSA P25074 oryza sa	_HUMAN P04632 homo s	_HUMAN P14651 homo	_NICSY P27484 nicotian	PETHY P09789 pet	PANTR 097775 pan	MOUSE P54728	Q03828 home	DROME Q27294 drosophil	MOUSE P43029 mus m	PIG P04574 sus s	HUMAN P22087 homo sapi	ARATH Q38896 arabi	Q00577 homo sa	K 042131 gallu	AOUSE P42669 mus mus	MAIZE P2902	MOUSE P70315 mus m	HCMVA P16768 humar	HUMAN P23490 homo	MOUSE P18165 mus m	_HCMVA P17151	BACSU P10946 bacill	896	81	5343 homo sap	PYCA CHICK CAILE AND
	g	! !																											1 I				<u></u>
		303	919		433	$\omega$		268		214	œ	Ч		7	0	S	9	O	201	7	7	2	280	2	9		æ	$\alpha$	26	9	9	332	440
do	Query	30.3	ف	4	4	4.	3.	ж Э	Э.	9	ж •	ъ.	ж Э	щ Э	ж Э	ς.	ä	ά.	ď.	7	2	ς.	ς.	ä	2	2	ς.	2	22.2	2	8	ς.	c
	ы	165	•	13	•	m	3	$^{\circ}$	$^{\circ}$	2	12	•	~	12	•	$^{\circ}$	12	•	$\sim$	7	$\sim$	•	~	12	21.	21.	•	21.	12	20.	20.	•	20
	Result No.	-	7	٣	4	Ŋ	9	7	œ	6	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	56	27	28	59	30	31	32

DOMAIN

P20265 homo sapien P31360 mus musculu O95409 homo sapien P10495 phaseolus v Q9h461 homo sapien O70133 mus musculu P26687 mus musculu P31361 mus musculu Q63262 rattus norv P20264 homo sapien Q03297 drosophila P13135 bos taurus
OC3N_HUMAN OC3N_MOUSE ZIC2_HUMAN GRP1_PHAVU FZD8_HUMAN DDX9_MOUSE TWST_MOUSE BRN1_MOUSE BRN1_RAT BRN1_RAT BRN1_RAT CANS_BOVIN
ненныннанн
443 445 533 252 694 1380 206 495 497 500 1093
22.0 22.0 22.0 21.9 21.9 21.8 21.8 21.8 21.8
120 120 120 119.5 119.5 119 119 119 119 119
88888888888888888888888888888888888888

## ALIGNMENTS

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                CHIB_POPTR STANDARD; PRT; 303 AA.

CHIB_POPTR STANDARD; PRT; 303 AA.

P29031;
01-DEC-1992 (Rel. 24, Created)
01-EEB-1996 (Rel. 33, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Acidic endochitinase WIN6.2B precursor (EC 3.2.1.14).
Acidic endochitinase WIN6.2B precursor (EC 3.2.1.14).
Populus trichocarpa (Western balsam poplar).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids I; Malpighiales; Salicaceae; Populus.
                                                                                                                                                                                                                                                                                                       MEDLINE=92003678; PubMed=1912489; Davis J.M., Clarke H.R.G., Bradshaw H.D. Jr., Gordon M.P.; Davis J.M., Clarke H.R.G., Bradshaw H.D. Jr., Gordon M.P.; "Populus chitinase genes: structure, organization, and similarity of translated sequences to herbaceous plant chitinases."; Plant Mol. Biol. 17:631-639(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Pfam; PFUUlos; Chirinbind.

PRINTS; PR00451; CHITINBINDNG.

ProDom; PD000574; Glyco_hydro_19; 1.

ProDom; PD000609; Chitin_bind; 1.

SMART; SM00270; ChtBD1; 1.

PROSITE; PS00026; CHITIN_BINDING; 1.

PROSITE; PS00773; CHITINASE_19_1; 1.

PROSITE; PS00774; CHITINASE_19_2; 1.

PROSITE; PS00774; CHITINASE_19_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIMILAR TO N-TERMINAL OF GLYCOSYL
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Ż
                                                                                                                                                                                                                                                                                                                                                                                                                                                         of
                                                                                                                                                                                                                                                                                                                                                                                                                               CONTAINING FUNGAL PATHOGENS.

-!- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages acetyl-D-glucosamine polymers of chitin.

-!- INDUCTION: By wounding.

-!- SIMILARITY: BELONGS TO CHITINASE CLASS IA WHICH IS SIMII CLASS IB IN THE CATALYTIC PART BUT WHICH INCLUDE A N-TERCHITIN-BINDING LECTIN DOMAIN (BELONGS TO FAMILY 19 OF GI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          BY SIMILARITY.
ACIDIC ENDOCHITINASE WIN6.2B.
CHITIN-BINDING (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            InterPro; IPR001002; Chitin_bind.
InterPro; IPR000726; Glyco_hydro_19.
Pfam; PF00187; chitin_binding; 1.
Pfam; PF00182; Glyco_hydro_19; 2.
PRINTS; PR00451; CHITINBINDNG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EMBL; X59995; CAA42612.1;
HSSP; P02877; 1HEV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            21
303
61
                                                                                                                                                                                                                                                                                               Ä.
                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A MEDLINE=92003678;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HYDROLASES)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SIGNAL
T 1
POPTR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HSSP
RESULT
                  CHIB
```

```
5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       androgen
                                                                                                                                 Gaps
                                                                                                                                                                                   63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               hormone
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        mutation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 R.E.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  acid sequences
                                                                                                                                                         FCRGG
                                                                                                                                                                                   -CDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         acid
                                                                                                                                                                                                                                                                                                                                                                                                        tomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         of
                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      the human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Lubahn D.B., Joseph D.R., Sar M., Tan J., Higgs H.N., Larson French F.S., Wilson E.M.; The human androgen receptor: complementary deoxyribonucleic cloning, sequence analysis and gene expression in prostate." Mol. Endocrinol. 2:1265-1275(1988).
                                                                                                                                                        --CGGGAVACQNYRQ
                                                                                                                                                                                                                                                                                                                                                                                                    Vertebrata; Euteleos; Hominidae; Homo.
                                                                                                      303;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A., AND VARIANT CAIS MET-866.
MEDLINE=90083302; PubMed=2594783;
Lubahn D.B., Brown T.R., Simental J.A., Higgs H.N., Migeon (Wilson E.M., French F.S.;
"Sequence of the intron/exon junctions of the coding region human androgen receptor gene and identification of a point nin a family with complete androgen insensitivity.";
Proc. Natl. Acad. Sci. U.S.A. 86:9534-9538(1989).
                                                                                                                                                                        LFLSLSVRGSAEQCGQQAGDALCPG-GLCCSSYGWCGTTADYCGDG---CQSQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ding domain is essential for human androgen receptor.";
                                                                                                                                Indels
                                                                                                         Length
                                                                              CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEDLINE=89112208; PubMed=3216866;
Lubahn D.B., Joseph D.R., Sar M., Tan J., Higgs H.N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=89098909; PubMed=2911578;
Tilley W.D., Marcelli M., Wilson J.D., McPhaul M.J.;
"Characterization and expression of a cDNA encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   hang C., Kokontis J., Liao S.;
Structural analysis of complementary DNA and amino
                                                                                                                                                                                                                                                                                                                      01-MAR-1989 (Rel. 10, Created)
01-APR-1990 (Rel. 14, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Androgen receptor (Dihydrotestosterone receptor)
AR OR NR3C4 OR DHTR.
                                                                              0E0D011DFA5CC8B8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            human and rat androgen receptors.";
Proc. Natl. Acad. Sci. U.S.A. 85:7211-7215(1988)
      (SPACER)
                                                                                                                  3e-08;
                                                                                                                                                                                                         GGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSKI
                                                                                                                                                                                                                      GGGGGGGGGGGGGGGG----DGYLSDIIPESMFDDML
                                                                                                         DB
                          SIMILARITY.
SIMILARITY.
SIMILARITY.
SIMILARITY.
                                                                                                                                Mismatches
                                                                                                                                                       LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG
                                                                                                                                                                                                                                                                                                                                                                                    Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Va
Mammalia; Eutheria; Primates; Catarrhini;
NCBI_TaxID=9606;
                                                                                                                   No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Govindan M.V.; "Specific region in hormone binding binding and trans-activation by huma Mol. Endocrinol. 4:417-427(1990).
                                                                                                        Score
                                                                                                                   Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=90258935; PubMed=2342476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TISSUE=Prostate;
MEDLINE=89017168; PubMed=3174628;
                                                    BY
BY
                                          \mathbf{B}\mathbf{X}
                                                                                                                               8;
81
303
39
45
852
860
832208 · MW;
                                                                                                      30.3%;
42.4%;
                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                 STANDARD;
                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Prostate;
MEDLINE=89098909;
62
82
24
33
38
56
                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EQUENCE FROM
                                                                                                                            42;
                                                                                                                                                                                                                                                                                                 ANDR_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 receptor."
                         DISULFID
DISULFID
DISULFID
DISULFID
SEQUENCE
                                                                                                      Query Match
                                                                                                                 Local
                                                                                                                                                                                                                                                                                                            P10275;
                                                                                                                                                                                                                                                                                   ANDR_HUMAN
                                                                                                                                                                                                                                 64
                                                                                                                                                                                                        59
                                                                                                                                                       4
                                                                                                                             Matches
                                                                                                                                                                                                                                                                        RESULT
                                      FT
FT
SO
  FT
FT
FT
                                                                                                                                                                                a
                                                                                                                                                                                                                                 qq
                                                                                                                                                                                                                                                                                                                                                ŏ
                                                                                                                                                       ō
                                                                                                                                                                                                                                                                                                                      TO
```

```
identification of mutations that cause androgen resistance: premature termination of the receptor protein at amino acid residue 588 causes complete androgen resistance.";
Mol. Endocrinol, 4:1105-1116(1990).
                                                                                                                                                                                                                                                                                                                                                                            Voorhorst M.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 POLYMORPHISM OF POLY-GLN REGION.
MEDLINE=97250535; PubMed=9096391;
Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Brufsky A., Talcott J., Hennekens C.H., Kantoff P.W.;
"The CAG repeat within the androgen receptor gene and its relationship to prostate cancer.";
Proc. Natl. Acad. Sci. U.S.A. 94:3320-3323(1997).
                                                                                                                                                                                                                                                                                                                                                               der Korput J.A.G.M.,
                                                                                 J.E., Wilson J.D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         gene
                                                                                                                                                                                                                                                                                                                                                                                                            structure and expression of a cDNA encoding the human
                                                                                                                                                                                                                                                              encoding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   the androgen receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Dahl D.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    INTERACTION WITH RAN.
MEDLINE=99329028; PubMed=10400640;
Hsiao P.-W., Lin D.-L., Nakao R., Chang C.;
"The linkage of Kennedy's neuron disease to ARA24, the first identified androgen receptor polyglutamine region-associated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       A.O., Trapman J.;
androgen receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Alvarado
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mhatre
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Giovannucci E., Stampfer M.J., Krithivas K., Brown M., Brufsky A., Talcott J., Hennekens C.H., Kantoff P.W.; Proc. Natl. Acad. Sci. U.S.A. 94:8272-8272(1997).
                                                                                                                                                                                                                                                                complementary DNA
                                                                                                                                                                                                                                                                                                                                                                              A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REVIEW ON VARIANTS AIS.
MEDLINE=93339360; PubMed=8339746;
Brown T.R., Scherer P.A., Chang Y.-T., Migeon C.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Beitel L.K.,
                                                                                                                                                                                                                                                                                                                                                                                                                                           Res. Commun. 153:241-248(1988)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lumbroso R.,
                                                                                                                                                                                                                                                                                                                                                      Kuiper G.G.J.M., van de
                                                                                Griffin
 86:327-331(1989)
                                                                          Marcelli M., Tilley W.D., Wilson C.M., Grif
McPhaul M.J.;
"Definition of the human androgen receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLYMORPHISM OF POLY-GLN REGION.
MEDLINE=92220629; PubMed=1561105;
Sleddens H.F.B.M., Oostra B.A., Brinkmann
"Trinucleotide repeat polymorphism in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Androgen resistance due to mutation of Clin. Invest. Med. 15:456-472(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20:1427-1427(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Chem. 274:20229-20234(1999)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Μ.,
                                                                                                                                                                                                                                            Chang C., Kokontis J., Liao S.; "Molecular cloning of human and rat
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REVIEW ON VARIANTS.
MEDLINE=93092459; PubMed=1458719;
Pinsky L., Trifiro M.A., Kaufman M.,
Kazemi-Esfarjani P., Sabbaghian N.,
Vasiliou M., Gottlieb B.;
                             SEQUENCE FROM N.A.
TISSUE=Prostate;
MEDLINE=91155943; PubMed=2293020;
                                                                                                                                                                                                           SEQUENCE OF 189-919 FROM N.A. MEDLINE=88178111; PubMed=3353726;
                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 468-919 FROM N.A. MEDLINE=88240407; PubMed=3377788;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      POLYMORPHISM OF POLY-GLY REGION.
                                                                                                                                                                                                                                                                                                                                                         Trapman J., Klaassen P., Kuipe
Faber P.W., van Rooij H.C.J.,
Mulder E., Brinkmann A.O.;
U.S.A.
                                                                                                                                                                                                                                                                             androgen receptors.";
Science 240:324-326(1988).
Sci.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Lu J., Danielsen M.;
Submitted (FEB-1995)
                                                                                                                                                                                                                                                                                                                                                                                                                          androgen receptor."
Biochem. Biophys. R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Nucleic Acids Res.
Acad.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TISSUE=Blood;
Lu J., Danielsen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         coactivator."
 Natl.
                                                                                                                                                                                                                                                                                                                                                                                                            "Cloning,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Biol.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ERRATUM.
Proc.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (AR)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                [10]
                [9]
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Ġ
```

Ghirri P.,

rsp.

```
qualitative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              "Androgen resistance associated with a mutation of the androgen receptor at amino acid 772 (Arg-->Cys) results from a combination of decreased messenger ribonucleic acid levels and impairment of receptor function.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          androgen receptor gene
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          androgen
                                                                                                                                                                                                                                                                                                                                         J.A.G.M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      receptor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                ky L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           C.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               D.,
                                                                                                                                 ity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     pu
                                                                                                                                                                                                                                                                                                                                                                                                            [19]
REVIEW ON VARIANTS.
MEDLINE=97169385; PubMed=9016528;
Gottlieb B., Trifiro M.A., Lumbroso R., Vasiliou D.M., Pinsl
Gottlieb B., Trifiro M.A., Lumbroso R., Vasiliou D.M., Pinsl
"The androgen receptor gene mutations database.";
"The androgen receptor 158-162(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               WARIANT LNCAP ALA-877.

MEDLINE-91083633; PubMed=2260966;

Veldscholte J., Ris-Stalpers C., Kuiper G.G.J.M., Jenster G
Berrevoets C.A., Claassen E., van Rooij H.C.J., Trapman J.,
Brinkmann A.O., Mulder E.;
"A mutation in the ligand binding domain of the androgen rehuman LNCaP cells affects steroid binding characteristics a response to anti-androgens.";
Biochem. Biophys. Res. Commun. 173:534-540(1990).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Functional characterization of naturally occurring mutant receptors from subjects with complete androgen insensitivit Mol. Endocrinol. 4:1759-1772(1990).
                                                                                                                  "Mutations of androgen receptor gene in androgen insensitives vndromes ".
                                                                                                                                                                                                                                                                                                                                         der Korput
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              VARIANT CYS-774.
MEDLINE-91310758; PubMed-1856263;
Marcelli M., Tilley W.D., Zoppi S., Griffin J.E., Wilson J. McPhaul M.J.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         patient with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      VARIANTS CAIS CYS-774; GLN-831 AND MET-866.
MEDLINE-91186983; PubMed=2082179;
Brown T.R., Lubahn D.B., Wilson E.M., French F.S., Migeon
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a family with
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Griffin J.E., Wilson
                                                                                             .
:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Griffin J.E.,
         androgen insensitivity 2:S62-S69(1993).
                                                                                              Boudon
                                                                                                                                                                                                                                                                                                                                         van
                                                                                                                                                                                                                                    B., Pinsky
database."
                                                                                                                                                                Steroid Biochem. Mol. Biol. 46:519-530(1993)
                                                                                                                                                                                                                                                                                                                                                                                         53:443-448(1995)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        in the DNA-binding domain of the ete testicular feminization in a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Septor function.";
Clin. Endocrinol. Metab. 73:318-325(1991)
                                                                                                                                                                                                                                                                                                                                   Brinkmann A.O., Jenster G., Ris-Stalpers C., Bruggenwirth H.T., Boehmer A.L.M., Trapman J. Androgen receptor mutations.";
J. Steroid Biochem. Mol. Biol. 53:443-448(1999)
                                                                                            .;
:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ļn
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         causes complete testicular feminization receptor-positive androgen resistance."; J. Clin. Invest. 87:1123-1126(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VARIANT PAIS CYS-763.
MEDLINE=91185626; PubMed=2010552;
McPhaul M.J., Marcelli M., Tilley W.D.,
Isidro-Gutierrez R.F., Wilson J.D.;
"Molecular basis of androgen resistance
                                                     REVIEW ON VARIANTS.
MEDLINE=94059770; PubMed=8240973;
Sultan C., Lumbroso S., Poujol N., Belon
                                                                                                                                                                                                                MEDLINE-95023089; PubMed=7937057;
Patterson M.N., Hughes I.A., Gottlieb
"The androgen receptor gene mutations
Nucleic Acids Res. 22:3560-3562(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Grino P.B.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     VARIANT CAIS PRO-617.
MEDLINE=91154385; PubMed=1999491;
                                                                                                                                                                                                                                                                                                                      PubMed=7626493;
          of human
                           Suppl.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Zoppi S.,
          ecular genetics o
J. Pediatr. 152
                                                                                                                                                                                                                                                                                                       REVIEW ON VARIANTS MEDLINE=95352489;
                                                                                                                                                                                                  REVIEW ON VARIANTS MEDLINE-95023089;
                                                                                              Sultan C., Lumbr
Lobaccaro J.-M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Marcelli M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             J.L.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         mutation
"Molecular
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                decreased r
receptor fo
                                                                                                                                                   syndromes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           McPhaul
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Corfen
                                                                                                                                                                                                                                                                                                                                                                                                J. S
[19]
                                                                                                                                                                                                                                                                                       [18]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RN
RP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RA
RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RA
RT
RT
```

```
Hum. Mol. Genet. 7:415-422(1998).

Hum. Mol. Genet. 6 a growth regulator and sometosensory information, as well as in face and body structure formation. May also have a role in heart development.

Hum. Mol. Genet. Genet. Genet. 6 and genet. Also surchial arch, highly expressed in striated muscle followed by liver, kidney, testis, brain, heart, lung and spleen.

Hum. Mol. Genet. 6 and Genet. 6 and genet. 6 and genet. 6 by liver, has all processes, eyelid, midbrain, medulla otic region, maxillary and mandibular components of the first branchial arch, nasal processes, eyelid, midbrain, medulla oblongata, limbs, dorsal root ganglia and genital tubercle. Also expressed in non-neuronal structures around the oral cavity and in hip and shoulder regions and in mesenchyme surrounding the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             specifying
                                                                                                                                                                                                                                                                                                                    442
                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 most conserved
development in
                                                                                                                                                                                                                                                                                                                 KLENPLDYGSAWAAAA----OCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SHX2_MOUSE STANDARD; PRT; 331 AA.
P70390; P70369;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
Short stature homeobox protein 2 (Homeobox protein Og12x) (OG-12)
(Paired family homeodomain protein Prx3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mu
           to high-dos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  van der Kleij A.A.M., Asoh S., Kozak C.A., Nirenberg M.W.,
Burbach J.P.H.;
"Homeobox gene Prx3 expression in rodent brain and extraneural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            enes.";
                                                                                                                                                                                                       25;
                                                                                                                                                 919;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .
Σ
                                                                                                                                                                                                                                                             -- AVACQNYRQF-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE=Embryo;
MEDLINE=98058757; PubMed=9371788;
van Schaick H.S.A., Smidt M.P., Rovescalli A.C., Luijten M van Schaick H.S.A., Smidt M.P., Rozak C.A., Nirenberg M.W., van der Kleij A.A.M., Asoh S., Kozak C.A., nirenberg M.W.,
                                                                                                                                                 Length
                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE OF 116-331 AND 235-331 FROM N.A. (ISOFORMS 1 STRAIN=BALB/c; TISSUE=Liver, and Embryo; MEDLINE=97008065; PubMed=8855241; Rovescalli A.C., Asoh S., Nirenberg M.W.; "Cloning and characterization of four murine homeobox Proc. Natl. Acad. Sci. U.S.A. 93:10691-10696(1996).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of the
heart d
              and responsive
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               94:12993-12998(1997)
                                                                                                                                 Score 142.5; DB 1;
Pred. No. 7.7e-06;
                                                                                                                                                                                                                                                                -GTIGCGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   member
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 during
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DEVELOPMENTAL EXPRESSION.

MEDLINE=98133920; PubMed=9466998;
Semina E.V., Reiter R.S., Murray J.C.;
"A new human homeobox gene OG12X is a language of the second of the se
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5
               androgen receptor
abnormality of the amero, androgen therapy.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A. (ISOFORMS 1 AND
                                                                                                                                                                                                                                                                                                                                                                                                        443 QLYGPCGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                               --CRGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                12 ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                          3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Sci. U.S.A.
                                                                                                                                                26.1%;
illarity 42.5%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Eukaryota; Metazoa;
                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Acad.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        musculus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         tissues.";
Proc. Natl.
                                                                                                                                                                                                          37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mammalia;
                                                                                                                                     Query Ma. -+ Local
                                                                                                                                                           Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          mouse."
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOUSE
                                                                                                                                                                                                                                                                                                                        386
                                                                                                                                                                                                                                                                                                                                                                               55
                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     --
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               <u>+</u>.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            R RT R
                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                     g
                                                                                                                                                                                                                                                                                                                                                                                    δ
```

in

rsp

<del>ر</del> ا

us-09-894-030

4

```
MEDLINE=95196953;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                129
154
191
113
119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    182
216
330
342
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       DNA_BIND
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ry Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best Loc
Matches
                                                                                                           3
        Κ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     QΫ
                                                                                                                                              way
                                                                                   collaboration
                                                                                                                                                                  commercial
                                                                                                                                                             for commercial
b.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2
                                                                                                     EMBL outstation trictions on it
                                                                                                                                               ou
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             107
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            stem and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                in
        TEINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -GGGGGGGGGGGGGGGGRSPVRELDMGAAERSRE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  comi; Mus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SMART; SMUUJOS; DOA, 1.

PROSITE; PS00027; HOMEOBOX_1; 1.

PROSITE; PS50071; HOMEOBOX_2; 1.

PROSITE; PS50803; OAR; 1.

HOMEODOX; DNA-binding; Developmental protein; Nuclear protein; Alternative splicing.

DNA_BIND 140 199 HOMEOBOX.

DOMAIN 313 326 OAR DOMAIN.

DOMAIN 59 82 POLY-GLY.

VARSPLIC 235 246 MISSING (IN ISOFORM 2).

SEQUENCE 331 AA; 34905 MW; D24D297E1D73A025 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14;
                                                                                                                                            S
                                                                      This SWISS-PROT entry is copyright. It is produced through a between the Swiss Institute of Bioinformatics and the EMB the European Bioinformatics Institute. There are no restriuse by non-profit institutions as long as its content imposified and this statement is not removed. Usage by and entities requires a license agreement (See http://www.isb-sicor send an email to license@isb-sib.ch).
  SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX PRC"BICOID" SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Craniata; Vertebrata; Euteleos'
Sciurognathi; Muridae; Murinae
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=92258392; PubMed=1582411; Sham M.H., Hunt P., Nonchev S., Papalopulu N., Graham A., Sham M.H., Hunt P., Nonchev S., Papalopulu N., Graham A., Boncinelli E., Krumlauf R.; "Analysis of the murine Hox-2.7 gene: conserved alternative transcripts with differential distributions in the nervous sthe potential for shared regulatory regions."; EMBO J. 11:1825-1836(1992).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 331;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ; DB 1; Land 1.6e-05; Ches 19;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HXB3_MOUSE STANDARD; PRT; 433 AA. P09026; P10285; Q61680; O1-NOV-1988 (Rel. 09, Created) O1-OCT-1996 (Rel. 34, Last sequence update) 15-JUL-1999 (Rel. 38, Last annotation update) Homeobox protein Hox-B3 (Hox-2.7) (MH-23). HOXB3 OR HOXB-3 OR HOX-2.7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ed. No. 1.6e
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 135;
                                    SIMILARITY: CONTAINS 1 OAR DOMAIN.
                                                                                                                                                                                                                    EMBL; U66918; AAC52833.1; -.
EMBL; U67055; AAC52834.1; -.
EMBL; U65071; -; NOT_ANNOTATED_CDS.
EMBL; U65072; AAC52831.1; ALT_INIT.
EMBL; U65072; AAC52832.1; ALT_INIT.
HSSP; P02836; 1ENH.
TRANSFAC; T03310; -.
MGD; MGI:1201673; Shox2.
InterPro; IPR001356; Homeobox.
InterPro; IPR001356; Homeobox.
InterPro; IPR003654; OAR_domain.
Pfam: PF00046; homeobox; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Rodentia;
NCBI_TaxID=10090;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      24.8%;
ilarity 38.9%;
Conservative 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRINTS; PR00024; HOMEOBOX. SMART; SM00389; HOX; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 D----SKITPQ 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       108 PGSPRLTEVSPE 119
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        57 AAGGGGGAGGGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
28; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         [2]
SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HXB3_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             94
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
STITIES SELECTION OF THE STATE OF THE SELECTION OF THE SE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    à
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Qγ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                           N [4]

P SEQUENCE OF 181-265 FROM N.A.

X MEDLINE=89091992; PubMed=2463210;
A Graham A., Papalopulu N., Lorimer J., McVey J.H., Tuddenham E.G.D.,
A Krumlauf R.;
Characterization of a murine homeo box gene, Hox-2.6, related to the Drosophila Deformed gene.";
Genes Dev. 2:1424-1438(1988).
C -1- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.
C -1- SUBCELLULAR LOCATION: Nuclear.
C -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEOBOX PROTEINS.
C -1- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEOBOX PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -> APTGRPPCRAVRCMWAG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RLWWWRPAVVAAAAVRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                  and its intron
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HSSP; P02833; 1SAN.

HSSP; P02833; 1SAN.

TRANSFAC; T01724; -.

MGD; MGI:96184; Hoxb3.

InterPro; IPR001827; Antennapedia.

InterPro; IPR001356; Homeobox.

Pfam; PF00046; homeobox; 1.

PRINTS; PR00024; HOMEOBOX.

RARRT; SM0389; HOX; 1.

PROSITE; PS00027; HOMEOBOX_1; 1.

PROSITE; PS00032; ANTENNAPEDIA; 1.

PROSITE; PS50071; HOMEOBOX_2; 1.

Homeobox; DNA-binding; Developmental protein; Nuclear protein;

Transcription regulation.
                                                                                                                                                                                                                      and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         433;
                                                                                                                                       SEQUENCE OF 152-361 FROM N.A.
MEDLINE=88054465; PubMed=2890503;
Lonai P., Arman E., Czosnek H., Ruddle F.H., Blatt C.;
"New murine homeoboxes: structure, chromosomal assignment, differential expression in adult erythropoiesis.";
DNA 6:409-418(1987).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GGGGGGGGGMSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                Brown W.M., Taylor G.R.; "The 5'-sequence of the murine Hox-b3 (Hox-2.7) gene contain multiple transcription-regulatory elements."; Int. J. Biochem. 26:1403-1409(1994).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ٨
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LC -> FV (IN REF. 3).
S -> L (IN REF. 3).
GAYGTPTMQGSPVYVGGGGY.
VAT (IN RFF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ANTP-TYPE HEXAPEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     G -> C (IN REF. 1).
A -> S (IN REF. 1).
GCGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 132.5; DB 1;
Pred. No. 3.3e-05;
; Mismatches 6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             T (IN REF. 3).
9AD3C922663612A6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             N (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (IN REF.
D -> N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GLY-RICH.
HOMEOBOX.
PubMed=7890121;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; X66177; CAA46951.1; -.
EMBL; U02278; AAB60496.1; -.
EMBL; M18168; AAA37840.1; -.
PIR; S20963; S20963.
PIR; C29585; C29585.
HSSP; P02833; 1SAN.
TRANSFAC; T01724; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    44353 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24.3%;
56.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     181
250
113
119
169
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              182
217
330
361
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      433 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
nes 30; Conser
```

ω,

.rsp

d

```
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                                                                                                                                                               ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d for commercial sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (shown here) and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    specifying or correction,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   as
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           icing.
e, liver,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           There are no restrictions ng as its content is in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              onserved
ment in
                                                                                                                                                                                                                                                                                                                                                                                                            ofacial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               development
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ROTEINS
                                                                                                                                                                               rotein
                                                                                                                                                                                                                                                     stomi;
                                                                                                                                                                                                                                                                                                                                                                             SHX2_HUMAN STANDARD; firs, SHX2_HUMAN STANDARD; c60967; c60902; 060903; 060465; 060467; c60902; 060903; 060465; 060467; c60902; 060903; 060465; 060467; created)

I 16-OCT-2001 (Rel. 40, Last annotation update)

T 16-OCT-2001 (Rel. 40, Last annotation update)

E Short stature homeobox protein 2 (Paired-related homeobox protein Stort) (Homeobox protein 0g12x).

SHOT) (Homeobox protein 0g12x).

SHOX2 OR SHOT OR 0G12X.

Homo sapiens (Human).

C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostom

C Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                              Rao
  183
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mol. Genet. 7:415-422(1998).

FUNCTION: May be a growth regulator and have a role in special systems involved in processing somatosensory infor as well as in face and body structure formation.

SUBCELLULAR LOCATION: Nuclear.

ALTERNATIVE PRODUCTS: 2 isoforms; 1/SHOX2A/SHOTA (shown by 2/SHOX2B/SHOTB/OG12XB; are produced by alternative splicitissue SPECIFICITY: Expressed in heart, skeletal muscle, lung, bone marrow fibroblast, pancreas and placenta.

DEVELOPMENTAL STAGE: Expressed during cranofacial develop
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ಥ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 and
                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
TISSUE=Fibroblast;
MEDLINE=98151525; PubMed=9482898;
Blaschke R.J., Monaghan A.P., Schiller S., Schechinger B., Padilla-Nash H., Ried T., Rappold G.A.;
"SHOT, a SHOX-related homeobox gene, is implicated in cranic brain, heart, and limb development.";
Proc. Natl. Acad. Sci. U.S.A. 95:2406-2411(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Eby the European Bioinformatics Institute. There are no restruse by non-profit institutions as long as its content modified and this statement is not removed. Usage by ancentitles requires a license agreement (See http://www.isb-sor send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  heart develops
   -ggegegegegessegegegegege
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      well as in heart.
SIMILARITY: BELONGS TO THE PAIRED FAMILY OF HOMEOBOX "BICOID" SUBFAMILY.
SIMILARITY: CONTAINS 1 OAR DOMAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the most
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (MAY-2001) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ^{\rm ot}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Semina E.V., Reiter R.S., Murray J.C.; "A new human homeobox gene OG12X is a member homeobox gene family and is expressed during mouse.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; AJ002367; CAA05341.1; ALT_INIT.
EMBL; AJ002368; CAA05342.1; ALT_INIT.
EMBL; BC008829; AAH08829.1; -.
EMBL; AF022654; AAC39662.1; ALT_INIT.
EMBL; AF023203; AAC39663.1; -.
HSSP; P06601; 1FJL.
MIM; 602504; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (ISOFORM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           IPR000047; HTH_repressr
IPR001356; Homeobox.
IPR003654; OAR_domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TISSUE=Craniofacial;
MEDLINE-98133920; PubMed=9466998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homeobox.
OAR_domain.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OF 116-331 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     homeobox;
      LKNSSPGTAEGCGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PF00046;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TISSUE=Muscle;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Strausberg R.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Submitted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro;
Pfam; PF0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
                                                                         HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Hum.
       142
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <del>-</del>:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <u>'</u>.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -
                                                         ESULT
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CELL WALL (POTENTIAL)
                                                                                                                                                                                                  (IN REF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                                                                                          ---GGGGGGGGGGGGGGGGGRSPVRELDMGAAER
                                                                                                                                                                                                                                                                                                                                                                                                          POTENTIAL.
GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.
GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                          16;
                                                                                                 protein;
                                                                                                                                                                                         E -> EGRRKPTKAEVQATLLLPGEAFRFL 2).

E -> D (IN REF. 1; CAA05341).

P -> S (IN REF. 2).

D -> N (IN REF. 2 AND 3)
                                                                                                                                                                                                                                                                                                                                       Length 331;
                                                                                                                                                                                                                                  E -> D (IN REF. 1; CAA05341).
P -> S (IN REF. 2).
D -> N (IN REF. 2 AND 3).
H -> L (IN REF. 3; AAC39663).
55431B073B3B2250 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Oryza sativa (Rice).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trach
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
NCBI_TaxID=4530;
                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                      SHOX2B)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lei M., Wu R.;
"A novel glycine-rich cell wall protein gene in rice.";
Plant Mol. Biol. 16:187-198(1991).
-!- FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WA
-!- SUBCELLULAR LOCATION: Cell wall (Potential).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    22, Created)
22, Last sequence update)
25, Last annotation update)
wall structural protein 1 precursor
                                                                                                    Nuclear
                                                                                                                                                                                                                                                                                                                                                                          19;
                                                                                                                                                                                      (IN ISOFORM
PRINTS; PR00024; HOMEOBOX.

PRINTS; PR00031; HTHREPRESSR.

SMART; SM00389; HOX; 1.

PROSITE; PS00027; HOMEOBOX_1; 1.

PROSITE; PS50071; HOMEOBOX_2; 2.

PROSITE; PS50803; OAR; 1.

Homeobox; DNA-binding; Developmental protein; Numerobox; DNA-binding; Developmental protein; Numerobox; DNA_BIND 140 199 HOMEOBOX.

DOMAIN 313 326 OAR.

DOMAIN 60 86 POLY-GLY.

VARSPLIC 235 246 MISSING (IN ISOFORM VARSPLICT 115 115 E.-> EGRRKPTKAEVQAT
                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                    132; bz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    165 AA
                                                                                                                                                                                                                                                                                                                                                          ed. No. 3e-(
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Repeat;
                                                                                                                                                                                                                                                                                                                                              Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          STRAIN-CV. INDICA-IR36;
MEDLINE-91370862; PubMed-1716496;
                                                                                                                                                                                                                                                                                                                                                            Pred
                                                                                                                                                                                                                                                                                                                                                                             12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Structural protein;
                                                                                                                                                                                                                                                                                                           34964 MW;
                                                                                                                                                                                                                                                                                                                                             24.2%;
37.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             X53596; CAA37665.1;
S13385; KNRZG1.
P30129; 4DPV.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 119
                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        159
                                                                                                                                                                                                                                        125
244
312
325
                                                                                                                                                                                                                                                                                                                                                                                                                                                  56 VRAAGGGGGGGGGG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SKITPO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SREPGSPRLTEVSPE
                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Rel. (Rel.
                                                                                                                                                                                                                                                                                                             AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               [1]
SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (Rel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24
                                                                                                                                                                                                                                    125
244
312
325
331 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                01-MAY-1992 (1
01-MAY-1992 (1
01-APR-1993 (1
Glycine-rich (GRP-1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        wall;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SKQD--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    GRP1_ORYSA P25074;
                                                                                                                                                                                                                                                        CONFLICT
CONFLICT
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PIR; S
HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       T 6
ORYSA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell
                                                                                                                                                                                                                                                                                                                                                                                                                  31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     91
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    105
```

```
2
                                                                                                                                                                                                                                                                                                                                                             subunit)
                                                                                                                                              of human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       MEDLINE=87066759; PubMed=3024120;
Miyake S., Emori Y., Suzuki K.;
"Gene organization of the small subunit of human calcium-activated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ests an
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       which
                                                                                                                               ---FCRGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            plasma
                                                                                                                                                                                                                                                                                                                                                                                                                       stomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Masumoto H
                                                                                                  .;
8
                                                                                                                                                                                                                                                                                                               13-AUG-1987 (Rel. 05, Created)
13-AUG-1987 (Rel. 05, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Calcium-dependent protease, small subunit (Calpain regulatory
(Calcium-activated neutral proteinase) (CANP).
CAPN4 OR CAPNS1 OR CAPNS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            "The crystal structure of calcium-free human m-calpain sugges
electrostatic switch mechanism for activation by calcium.";
Proc. Natl. Acad. Sci. U.S.A. 97:588-592(2000).
-! FUNCTION: Calcium-regulated non-lysosomal thiol-protease
catalyze limited proteolysis of substrates involved in
cytoskeletal remodelling and signal tranduction.
-! SUBUNIT: Heterodimer of a large (catalytic) and a small
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Burkhart-Schultz
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ohno S., Emori Y., Suzuki K.; "Nucleotide sequence of a cDNA coding for the small subunit calcium-dependent protease."; Nucleic Acids Res. 14:5559-5559(1986).
                                                                                                                                                                                                                                                                                                                                                                                                                       Euteleo
                                                                       Length 165;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
MEDLINE=20105516; PubMed=10639123;
Strobl S., Fernandez-Catalan C., Braun M., Huber R., Masum
Nakagawa K., Irie A., Sorimachi H., Bourenkow G., Bartunik
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             the
                                                                                                                                                                                                                                                                                                                                                                                                 Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                  Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 [3]
SEQUENCE FROM N.A.
Lamerdin J.E., McCready P.M., Adamson A.W., Burkhart-scu
Garcia E., Kyle A., Ramirez M., Stilwagen S., Garnes J.,
Garcia E., Kyle A., Quan G., Montgomery M., Ow D.,
Danganan L., Bruce R., Quan G., Montgomery M., Ow D.,
Kobayashi A., Olsen A.O., Carrano A.V.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ဌ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Strobl S., Fernandez-Catalan C., braun m., muxer ..., Nakagawa K., Irie A., Sorimachi H., Bourenkow G., Bartu Suzuki K., Bode W.; "The crystal structure of calcium-free human m-calpain "The crystal structure of calcium-free human by calcium."
(TYR-RICH).
(TYR-RICH).
(TYR-RICH).
E36CE31C3650AC9A CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Cytoplasmic; Translocates
                                                                ; DB 1; L.
2.6e-05;
-has 22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (MAY-2001) to the EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                        268 AA
                                                                                                  Mismatches
                                                                                                                             GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQ-
                                                                      Score 130;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   neutral protease.";
Nucleic Acids Res. 14:8805-8817(1986)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
MEDLINE-86286563; Pubmed=3016651;
R2
R2
R2
                                                                                                  3;
                                            MW.
                                                                       23.9%;
62
99
138
13536
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             (regulatory) subunit. SUBCELLULAR LOCATION:
                                                                                                  Conservative
                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             and Uterus;
                                                                                     Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FROM N.A.
                                            Æ
                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI TaxID=9606;
56
93
132
165 #
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FROM N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Strausberg R. Submitted (MA
                                                                                                                                                                                    16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TISSUE=Skin,
                                                                                                                                                                                                                                                                                   CANS_HUMAN
PO4632;
                                                                                                31;
                                                                                                                                                                                      73 GGGG
                                                                                                                                                                                                                 GGSG
REPEAT
REPEAT
REPEAT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                     Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  neutral
                                                                                                                                                         66
                                                                                                 Matches
                                                                                                                               18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1
                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                        g
FT
FT
SO
                                                                                                                                                                                                                d
                                                                                                                                                                                                                                                                                                                              DE DE GN
                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                      ğ
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              , Migliaccio
Boncinelli E
membrane upon Ca++ binding (By similarity).
SIMILARITY: Contains 4 EF-hand calcium-binding domains.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 268;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           17B87A8E47A90632 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    € 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       3D-structure.
GLY-RICH (HYDROPHOBIC).
CALCIUM-BINDING.
EF-HAND 1.
EF-HAND 2.
ANCESTRAL CALCIUM SITE 3
ANCESTRAL CALCIUM SITE 4
POLY-GLY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SITE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=90098876; PubMed=2574852;
Acampora D., D'Esposito M., Faiella A., Pannese M., Morelli F., Stornaiuolo A., Nigro V., Simeone A., Fine human HOX gene family.";
Nucleic Acids Res. 17:10385-10402(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 129; DB 1; I
Pred. No. 4.7e-05;
0; Mismatches 18;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     HXB3_HUMAN STANDARD; PRT; 431 AA. P14651; P17484; 095615; 01-APR-1990 (Rel. 14, Created) 01-APR-1990 (Rel. 14, Last sequence update) 01-MAR-2002 (Rel. 41, Last annotation update) HOMEODOX protein HOX-B3 (Hox-2G) (Hox-2.7).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          POLY-GLY.
POLY-PRO.
                                                                                                                                                                                                                                                                 EMBL; X04106; CAA27726.1; -.
EMBL; M31511; AAA35646.1; -.
EMBL; M31502; AAA35646.1; JOINED.
EMBL; M31504; AAA35646.1; JOINED.
EMBL; M31504; AAA35646.1; JOINED.
EMBL; M31506; AAA35646.1; JOINED.
EMBL; M31509; AAA35646.1; JOINED.
EMBL; M31509; AAA35646.1; JOINED.
EMBL; M31509; AAA35646.1; JOINED.
EMBL; M31509; AAA35646.1; JOINED.
EMBL; AD001527; AAB51183.1; -.
EMBL; AC002984; AAB81546.1; -.
EMBL; BC000592; AAH00592.1; -.
EMBL; BC007779; AAH07779.1; -.
EMBL; A23650; CIHUL.
PDB; 1KFU; 07-DEC-01.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        56
83
28316 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IPR002048; EF-hand.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ^{\circ}
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity 50.0%;
Matches 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           യ വ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       268 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Calcium-binding; F
DOMAIN 1
DOMAIN 96
CA_BIND 152
CA_BIND 182
DOMAIN 217
DOMAIN 247
DOMAIN 247
DOMAIN 35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9 KGGGGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1KFU; 0'
114170;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; PF0
PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  HXB3_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PDB;
MIM;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 8
  TTT W DD R R R B DD R R B DD R R B DD R DD R B DD R DD R
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ģ
```

-

Pag

rsp.

```
Matches
                                                  35
                                                                            146
                        Matches
                                                                                                                 SULT
                                                                                                                                                                                                           ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 its
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   e EMBL outstation restrictions on it
                                                                                                                                                                                                                                                                                                                                 MEDLINE=90215256; PubMed=2576652;

MEDLINE=90215256; PubMed=2576652;

MEDLINE=90215256; PubMed=2576652;

MEDLINE=90215256; PubMed=2576652;

MEDLINE=90215256; PubMed=2576652;

A Boncinelli E., Acampora D., Pannese M., D'Esposito M., Somma R.,

"Organization of human class I homeobox genes.";

Genome 31:745-756(1989).

C -!- FUNCTION: SEQUENCE-SPECIFIC TRANSCRIPTION FACTOR WHICH IS PART OF

A DEVELOPMENTAL REGULATORY SYSTEM THAT PROVIDES CELLS WITH

SPECIFIC POSITIONAL IDENTITIES ON THE ANTERIOR-POSTERIOR AXIS.

C -!- SUBCELLULAR LOCATION: Nuclear.

C -!- DEVELOPMENTAL STAGE: EXPRESSED IN WHOLE EMBRYOS AND FETUSES AT

C -!- SUBCELLULAR STAGE: EXPRESSED IN WHOLE EMBRYOS AND FETUSES AT

C -!- SIMILARITY: BELONGS TO THE ANTP FAMILY OF HOMEOBOX PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                 OF
              ٦.
                                                                                                                                                                                                                                                                                                                                                                                                                  PART
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AT
                                                                                                                                                .Y
ister.";
              Lawrence H
                                        ses
                                                                                                                                                                                                                                                                               erior
                                                                                                                    , · .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SM00389; HOX; 1.
PROSITE; PS00027; HOMEOBOX_1; 1.
PROSITE; PS00032; ANTENNAPEDIA; 1.
PROSITE; PS50071; HOMEOBOX_2; 1.
Homeobox; DNA-binding; Developmental protein; Nuclear protein; Transcription regulation.
                                                                                                                    Kidd K.K., Busygina V., DeMille M.M.C., Speed W.C., Ruggeri Kidd J.R., Pakstis A.J.; "Overall linkage disequilibrium in 33 populations for highly informative multisite haplotypes spanning the HOXB gene clus Am. J. Hum. Genet. 67:235-235(2000).
                                                                                                                                                                                                                             Pannese M.,
                                         can
                                                                                                                                                                                                                                                                           along the ant
                                                      and progr
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Elthe European Bioinformatics Institute. There are no restuse by non-profit institutions as long as its content
              Hugo P.,
                                        cells
                                                                               to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         . -> HV (IN REF. 2).
941706EDCC2975E5 CRC64;
                                                                                                                                                                                                                                                                                           system.";
                                                                                                                                                                                               TISSUE=Placenta;
MEDLINE=8937858; PubMed=2570724;
Giampaolo A., Acampora D., Zappavigna V., Pannese M.D'Esposito M., Care A., Faiella A., Stornaiuolo A., Simeone A., Boncinelli E., Peschle C.; "Differential expression of human HOX-2 genes along posterior axis in embryonic central nervous system." Differentiation 40:191-197(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANTP-TYPE HEXAPEPTIDE GLY-RICH. HOMEOBOX.
SEQUENCE FROM N.A.
Sauvageau G., Thorsteinsdottir U., Hough M.R., Hug
Largman C., Humphries R.K.;
"Deregulated expression of HoxB3 in hematopoietic
defective development of alpha beta T Lymphocytes
myeloproliferation.";
Submitted (MAY-1996) to the EMBL/GenBank/DDBJ data
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X16667; CAA34657.1; -.
EMBL; U59298; AAD10852.1; -.
EMBL; AF287967; AAG31555.1; -.
EMBL; X16175; CAA34297.1; -.
PIR; S07543; WJHU2G.
PIR; D37042; D37042.
HSSP; P02833; 1SAN.
TRANSFAC; T01723; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Pfam; PF00046; homeobox; 1. PRINTS; PR00025; ANTENNAPEDIA. PRINTS; PR00024; HOMEOBOX.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       44344 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134
178
247
200
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AA;
                                                                                                          SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  129
154
188
199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        431
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DNA_BIND
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                         [5]
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              3
                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          a putative glycine-rich
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nicotiana sylvestris (Wood tobacco).
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
Spermatophyta, Magnoliophyta, eudicotyledons, core eudicots,
Asteridae, euasterids I, Solanales, Solanaceae, Nicotiana.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ອວວວວວວວວວວວວ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         =
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               18;
                                                                               12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            wall
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Length 214;
                        Length 431;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <u>=</u>
=
=
=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  cell
                                                                                   Indels
                                                                                                                                                                                                          180
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (CSD)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Obokata J., Ohme M., Hayashida N.;
"Nucleotide sequence of a cDNA clone encoding a putat protein of 19.7 kDa in Nicotiana sylvestris.";
Plant Mol. Biol. 17:953-955(1991).
-!- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSI-: SIMILARITY: CONTAINS 2 C2HC-TYPE ZINC FINGERS.
-!- SIMILARITY: CONTAINS 2 CCHC-TYPE ZINC FINGERS.
-!- CAUTION: Was originally (Ref.1) thought to be a structural protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ಥ
                                                                                                                                                                                W; ELU-
SCORE 128; DB 1; Le
---A. NO. 4.8e-05;
Score 129; DB 1; Le
Pred. No. 6.8e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        E28DB84538F2A0AA
                                                                                                                                                                                                                                                                                                                                                                                                                      23, Created)
23, Last sequence update)
41, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCR
                                                                                                                                                                                                                                                                                                                                                                    AA
                                                   ed. No. 6.8e
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CCHC-TYPE
GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CCHC-TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 csD.
GSD.
GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF00313; CSD; 1.
Pfam; PF00098; zf-CCHC; 2.
PRINTS; PR00939; C2HCZNFINGER.
PRINTS; PR00050; COLDSHOCK.
ProDom; PD000621; Cold_shock; 1.
SMART; SM00357; CSP; 1.
SMART; SM00343; ZnF_C2HC; 2.
PROSITE; PS00352; COLD_SHOCK; 1.
PROSITE; PS50158; ZF_CCHC; 2.
RNA-binding; Repeat; Zinc-finger.
                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      PubMed=1912512;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; X60007; CAA42622.1; -.
PIR; S17731; KNNT2S.
HSSP; P15277; 1MJC.
InterPro; IPR002059; Cold_shock
InterPro; IPR001878; Znf_CCHC.
                                                                                 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          19746 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23.5%;
                         23.7%;
58.7%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Conservative
                                                                                      Conservative
                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   75
158
174
195
211
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Repeat;
8 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Glycine-rich protein GRP-2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel. 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FROM N.A.
                                                          Similarity 27; Conser
                                                                                                                                                                                     MEDLINE=92003709;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      82
157
176
194
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=4096;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               214
                                                                                                                                                                                                                                                                                                                                                                                                                             01-AUG-1992
01-AUG-1992
01-MAR-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               30;
                                                                                                                                                                                                                                                                                                                                                                    GRP2_NICSY
P27484;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ZN_FING
DOMAIN
ZN_FING
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE
                               Query Match
                                                            Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      18
```

.rsp

3

-030-

us-09-894

 $\infty$ 

```
34;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        R23B_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                   DNA_BIND
ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 12
R23B_MOUSE
ID R23B M
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       442
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     55
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 385
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ã
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                           This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                 L STRANDS.
TWO RELATED
                                                                                                                                                                                                                                         petunia.";
                                                                                                                                                                                                                                                            OTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      75
                                                                                                                                                        Petunia hybrida (Petunia).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
                                                                                                                                                                                                                                                                                       APABLE
 -GGGSGGGSGCFKCGESGHFARDCSQSGGGGGGGGRFGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     -GCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GGLGG
                                                                                                                                                                                                                                                                                                                      ABOUT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GLYCINE-RICH CELL WALL STRUCTURAL PROTEIN 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    24;
                                                                                                                                                                                                                                                                               Š
                                                                                                                                                                                                                          Condit C.M., Meagher R.B.;
"A gene encoding a novel glycine-rich structural protein of p Nature 323:178-181(1986).
-! FUNCTION: RESPONSIBLE FOR PLASTICITY OF THE CELL WALL (PC-!- SUBCELLULAR LOCATION: Cell Wall (Potential).
-! MISCELLANEOUS: THIS PROTEIN CONTAINS 67% GLYCINE RESIDUES -!- MISCELLANEOUS: 90% OF THE MATURE PROTEIN RESIDUES ARE CAFFORMING A BETA-PLEATED SHEET COMPOSED OF 8 ANTI-PARALLEL-!- MISCELLANEOUS: THE GLYCINE-RICH REGION IS COMPRISED OF TWEATLIES OF REPEATS, F1 AND F2, EACH REPEAT CONTAINING AE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              -6666666666666666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 384;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CRC64;
                                                                                                                                          precursor
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 127; DB 1; Le
Pred. No. 9.4e-05;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GLY-RICH.
C8541C549417D18C
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Signal.
                                                                                                         10, Created)
10, Last sequence update)
23, Last annotation update)
wall structural protein 1 p
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Created)
Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein; Repeat;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -AGGGVGGGVGGGAA -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                384 G
28777 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            23.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                          X04335; CAA27866.1; -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  GSGLGKAQCAALWLQCASGGTI
                                                                                       STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             STANDARD;
                                                                                                         01-MAR-1989 (Rel. 10,
01-MAR-1989 (Rel. 10,
01-AUG-1992 (Rel. 23,
Glycine-rich cell wall
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Structural
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               39,
39,
                                                                                                                                                                                                                                                                                                                                                                                                                                                4DPV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        384 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
                                                                                                                                                                                                                   FROM N.A
                                     198
                                                                                                                                                                                               NCBI_TaxID=4102;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (Rel
(Rel
(Rel
                   79
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    1
28
                                                                                                                                                                                                                                                                                                                                                                                                                                  A26099; A
; P30129;
wall; Str
                                     GGGGCYK
                   GGGGMSK
                                                                                                         01-MAR-1989
01-MAR-1989
01-AUG-1992
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ANDR_PANTR
097775;
30-MAY-2000
30-MAY-2000
GGGYG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      GGGFG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        76
                                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         G
                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP;
                                                                   T 10
PETHY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            G
                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 11
ANDR_PANTR
                                                                                                                                                GRP-1
145
                   73
                                     192
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      232
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9/
                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           270
                  \Omega Y
g
                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AC
DT
DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

EMBL, U94177; AAC73048.1; -.
HSSP; P06536; IGDC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                   collaboration -
                                                                                                                                                                                                                                                                                                                                          DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN, A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN. SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
                                                                                                                                                                                                                       disease.";
J. Mol. Evol. 47:334-342(1998).
-!- FUNCTION: THE STEROID HORMONES AND THEIR RECEPTORS ARE INVOLVED
THE REGULATION OF EUKARYOTIC GENE EXPRESSION AND AFFECT CELLULAR
PROLIFERATION AND DIFFERENTIATION IN TARGET TISSUES.
-!- SUBCELLULAR LOCATION: Nuclear.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                       SEQUENCE FROM N.A.
MEDLINE=98404153; PubMed=9732460;
Choong C.S., Kemppainen J.A., Wilson E.M.;
"Evolution of the primate androgen receptor: a structural basis for
                                                                Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KLENPLDYGSAWAAAA----QCRYGDLASLHGAGAAGPGSGSPSAAASSSWHTLFTAEEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HSSP; P06536; 1GDC.
InterPro; IPR001103; Androgen_recep.
InterPro; IPR001628; zf-C4.
InterPro; IPR001628; zf-C4.
InterPro; IPR001628; zf-C4.
Pfam; PF02166; Androgen_recep; 1.
Pfam; PF00104; hormone_rec; 1.
Pfam; PF00105; zf-C4; 1.
PRINTS; PR00047; STROIDFINGER.
SMART; SM00430; HOLI; 1.
SMART; SM00399; ZnF_C4; 1.
PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
Receptor; Transcription regulation; DNA-binding; Nuclear pro
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            --GTIGCGGG---AVACQNYRQF
                                                                                       Pan.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NUCLEAR RECEPTOR-TYPE,
C4-TYPE,
C4-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CRC64;
                                                               Vertebrata; F
L; Hominidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            126.5; DB
Androgen receptor (Dihydrotestosterone receptor) AR OR NR3C4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-GLY.
601B9BD4E697DAA4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          LIGAND-BINDING
POLY-GLN.
                                                               Craniata; Ve
Catarrhini;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ed. No. 0.0 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    POLY-GLN. POLY-GLN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-PRO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               POLY-ALA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            -CRGGGGGGGGGGGGGGGGMSKF
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               QLYGPCGGGGGGGGGGGGGGGGGAVAPY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ELENQKGSGLGKAQCAALWLQCASG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                   Pan troglodytes (Chimpanzee).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
NCBI_TaxID=9598;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        380 P
401 P
464 P
98402 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              23.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                616
571
611
911
78
88
                                                                                                                                                                                                                                                                                                                                                                                                         NR3 SUBFAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             551
587
682
682
57
1192
371
395
```

4;

416 AA.

STANDARD;

us-09-894-030-3.rsp

```
a collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    way
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d for commercial sib.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ouse
sion repair
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              I IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ij
                                                                                                                                                                                                                                                ae; Mus
                                                                                                                                                                                                                            Euteleostomi;
                                                                                                                          repair
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PART
URE I
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ;
0
P54728;
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001 (Rel. 40, Last annotation update)
UV excision repair protein RAD23 homolog B (MHR23B) (XP-C repair protein RAD23 homolog B (MHR23B).
RAD23B OR MHR23B.
Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostc Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
STRAIN=BALB/C; TISSUE=Testis;
MEDLINE=96403997; Pubmed=8808275;
wan der Spek P.J., Visser C.E., Hanaoka F., Smit B.,
Hagemeijer A., Bootsma D., Hoeijmakers J.H.J.;
"Cloning, comparative mapping, and RNA expression of the mouse homologues of the Saccharomyces cerevisiae nucleotide excision gene RAD23.";
Genomics 31:20-27(1996).
-!- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.
-!- SUBUNIT: HETERODIMER OF A 125 kDa SUBUNIT (P125) AND OF A 58 kDa SUBUNIT (P58).
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
-!- SIMILARITY: CONTAINS 2 UBA DOMAINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ē
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       416
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           is produced through
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         between the Swiss Institute of Bioinformatics and the I the European Bioinformatics Institute. There are no restuse by non-profit institutions as long as its content modified and this statement is not removed. Usage by an entities requires a license agreement (See http://www.isborsend an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              rmatics and the There are no rest
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              POLY-GLY.
13E0245A6D892205 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      375
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ·. 6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Last annotation update) homolog protein 2 (EVX-2)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 126; DB 1;
Pred. No. 0.00012
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57 GGGGGGGGGGGGGGGGGMSKFDDFDLDVVKVSKQDSK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   336 GGQGGGGGGGGGGGGGGGIAEAGSGHMNYIQVTPQEKE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SITE; PS50053; ÜBIQUITIN_2; 1.
damage; DNA repair; Nuclear protein.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               476
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This SWISS-PROT entry is copyright. It
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            POLY-ALA.
POLY-THR.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EVX2_HUMAN STANDARD; PRT; 4
Q03828;
01-OCT-1996 (Rel. 34, Created)
16-OCT-2001 (Rel. 40, Last sequence up
16-OCT-2001 (Rel. 40, Last annotation
Homeobox even-skipped homolog protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; X92411; CAA63146.1; -. HSSP; P54725; 1DV0.
MGD; MGI:105128; Rad23b.
InterPro; IPR000449; UBA.
InterPro; IPR000626; Ubiquitin.
Pfam; PF00627; UBA; 2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    43516 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 23.1%;
52.5%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PF00627; UBA; 2.
PF00240; ubiquitin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           261
270
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SMART; SM00165; UBA; 2
SMART; SM00213; UBQ; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity 21; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         255
262
336
416 A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PROSITE;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HSSP;
MGD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EVX2_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT 13
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DNA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Best
                    AC DIT DIT DE CONTROL 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     D'Esposito M., Morelli F., Acampora D., Migliaccio E., Simeone A., Boncinelli E.;
Boncinelli E.;
"EVX2, a human homeobox gene homologous to the even-skipped segmentation gene, is localized at the 5' end of HOX4 locus on chromosome 2.";
Genomics 10:43-50(1991).
-!- SUBCELLULAR LOCATION: Nuclear.
-!- DEVELOPMENTAL STAGE: EXPRESSED DURING EARLY EMBRYOGENESIS AND NEUROGENESIS IN A BIPHASIC MANNER.
-!- DEVELOPMENTY: BELONGS TO THE EVEN-SKIPPED FAMILY OF HOMEOBOX PROTEINS.
           Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CAZ_DROME STANDARD; PRT; 404 AA.
Q27294; Q24445; Q9VXI4;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
RNA-binding protein cabeza (Sarcoma-associated RNA-binding fly homolog) (P19).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 476;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Indels
                                                                         Birren B., Linton L., Nusbaum C., Lander E.;
Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases
           Craniata; Vertebrata; I
Catarrhini; Hominidae;
           Vertebrata;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  23;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  0.00014
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ed. No. 0.00
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 126;
                                                                                                                                                                                                                                                                                                                                                                                          AC009336; -; NOT_ANNOTATED_CDS.

M59983; AAA52414.1; -.

M59982; AAA52414.1; JOINED.

P14653; 1B72.

142991; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSAPCSCLSCHSSQSAAAAAAAAALGS--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred
                                                                                                             SEQUENCE OF 144-300 FROM N.A. MEDLINE=91257849; PubMed=1675198;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    23.1%;
illarity 46.7%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
ses 28; Conser
                                                               SEQUENCE FROM N.A.
                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                         EMBL;
EMBL;
HSSP;
MIM; 1
                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 14
CAZ_DROME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                383
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AC
DT
DT
DE
ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qi
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δŽ
```

-894-030-3.rsp

60-

us

```
Ketchum K.A.,
Lai Z.,
Lin X.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Jes F.,
Dietz S.M.,
V B.C., Dunn P.,
Fleischmann W.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Zheng L.,
Smith H.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Miklos G.L.G.,
S., Baldwin D.,
Beasley E.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          acleb J.M.,
V., Reese M.G.,
F., Shen H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein,
human sarcoma
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        nandra I.,
                                                                                         comorpha;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                er B.D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               J.D.,
R.F.,
Son S.N.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              A.,
n D.L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Q.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Wang X.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .
S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Adams M.D., Celniker S.E., Holt R.A., Evens C.A., Gocayne J.D., Adams M.D., Celniker S.E., Holt R.A., Evens C.A., Gocayne J.D., Adams M.D., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F. Gocoge R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.I. Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Raydayani A., An H. J., Andrews Prannkoch C., Baldwin B. Ballew R.M., Beason R.Y., Beanos P.V., Berman B.P., Bhandari D., Bolshakov S., Ballew R.M., Bacchan M.R., Bouck J., Baryaktaroglu L., Beasley E.J. Benchan M.R., Bouck J., Baryaktaroglu L., Beasley E.J. Benchan M.R., Bouck J., Baryaktaroglu L., Beasley E.J. Benchan M.R., Bouck J., Barkstein P., Bolshakov S., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chadr. Gerry J.M., Caadley S., Dahlke C., Davenport L.B., Davies P., Gerry J. M., Caadley S., Dahlke C., Davenport L.B., Davies P., Gerry J.M., Cabngelista C.C., Ferraz C., Ferriera S., Fleischman R.D., Barrey D., Heiman T.J., Hernandez J.R., Houck J., Houston K.A., Howland T.J., Wei M.-H., Ibegwam R., Jalali M., Kalush F., Karpen G.H., Kravitz S., Kulp D., Lai Z., Liang Y., Lin X., A Hostin D., Huston M. M., Moylan Y.C., McLeod M.P., Mortherson D.L. Markin B., Murphy L., Muzhy D.M., Nelson D.L. Markin B., Murphy L., Muzhy D.M., Nelson D.L. Markin G., Siden Flaman S., Pollard J., Pollard J., Moslon R., Shue B.C., Siden Flaman S., Pollard J., Pollard J., Moslon R., Shue B.C., Siden Flaman S., Pollard J., Weissenbach J., Wang Z.-Y. Wassarman D.A., Weinstock G.M., Weissenbach J., Wang Z.-Y. Wassarman D.A., Weinstock G.M., Weissenbach J., Wang Z.-Y., Wassarman D.A., Velter E., Wang G., Zhen R., Shen R., Shore 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        T.,
                                                           Hexapoda; Insecta
1; Brachycera; Mus
                                                                                                                                                                                                                                                                            Immanuel D., Zinszner H., Ron D.; "Associated RNA-binding fly hassociation of SARFH (sarcoma-associated RNA-binding fly haith regions of chromatin transcribed by RNA polymerase II. Mol. Cell. Biol. 15:4562-4571(1995).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                novel RNA binding genes involved in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (APR-1988) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Rubin G.M., Venter J.C.;
Drosophila melanogaster.";
                           Drosophila melanogaster (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; He
Pterygota; Neoptera; Endopterygota; Diptera;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Drosophila gene encoding a
                                                                                                                                                                                                       5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      23:835-843(1995)
                                                                                                                                                                                                   N.A. (ISOFORMS 1 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A. (ISOFORM 1).
STRAIN-BERKELEY;
MEDLINE=20196006; PubMed=10731132;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             shares homology with EWS and TLS,
                                                                                                                                                                                                                              STRAIN=CANTON-S;
MEDLINE=95349623; PubMed=7623847;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               STRAIN=CANTON-S;
MEDLINE=95223793; PubMed=7708500
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 F.W., Zhong W
E.W., Rubin (
ince of Drosopl
                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A. (ISOFORM 1)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE OF 212-261 FROM N.A. STRAIN=OREGON-R;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            sequence of Dros. :2185-2195(2000).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            39-404 FROM N.A
CG3606
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 formation.";
Nucleic Acids Res.
                                                                                                               Ephydroidea; Dro
NCBI_TaxID=7227;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STRAIN-OREGON-R;
 SARFH OR
                                                                                                                                                                                                     FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Stolow D.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genome
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  287
                                                                                                                                                                                                   SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    "Cabeza,
```

RETAR BARBARA BARBARA

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                      Haynes S.R., Rebbert M.L., Mozer B.A., Forquignon F., Dawid I.B.;

Then repeat sequences are GGN clusters and encode a glycine-rich domain in a Drosophila cDNA homologous to the rat helix destabilizing protein.";

Proc. Natl. Acad. Sci. U.S.A. 84:1819-1823(1987).

-!- FUNCTION: MAY PARTICIPATE IN A FUNCTION COMMON TO THE EXPRESSION OF MOST GENES TRANSCRIBED BY RNA POLYMERASE II.

-!- SUBCELLULAR LOCATION: Nuclear.

-!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);

C -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);

C -!- TISSUE SPECIFICITY: UBIQUITOUS. ENRICHED IN THE BRAIN AND CENTRAL NERVOUS SYSTEM DURING EMBRYOGENESIS. ENRICHED IN THE ADULT HEAD.

C -HEADS AND TORSOS) ONLY THE TYPE 2 ISOFORM IS DEFENDED.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EVKLEELEN--QKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGGGGG 64
                                                                                                                                                                                                                                                                                                                                                                                                                                    FROM THE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                              DEVELOPMENTAL STAGE: EXPRESSED IN THE DEVELOPING EMBRYO FROM THE EARLIEST STAGES OF CELLULARIZATION AND IS SUBSEQUENTLY FOUND IN MANY CELL TYPES.
MISCELLANEOUS: 'CABEZA' MEANS 'HEAD' IN SPANISH.
SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
SIMILARITY: CONTAINS 1 RANBP2-TYPE ZINC FINGER.
SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        - RGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          27;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         404;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GPMRNDGG -> MVDQEKRWS (IN REF 7062A0446BEA5984 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Metal-binding; RNA-binding;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AND 5).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MISSING (IN ISOFORM 1).

PNY ~> LFI (IN REF. 4).

P -> H (IN REF. 3).

G -> GG (IN REF. 3).

MISSING (IN REF. 3).

D -> E (IN REF. 3).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          GLY-KICH.
RNA-BINDING (RRM).
GLY-RICH.
RANBP2-TYPE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         -GGGGRGGFGGR - -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 125.5; DB Pred. No. 0.00013
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DGGPMRNDGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FALSE_NEG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pfam; Frous, zf-Ranbp; 1.

Pfam; PF00641; zf-Ranbp; 1.

SMART; SM00360; RRM; 1.

SMART; SM00547; ZnF_RBZ; 1.

PROSITE; PS50102; RRM; 1.

PROSITE; PS00030; RRM_RNP_1; FALE PROSITE; PS01358; ZF_RANBP2_1; 1

PROSITE; PS50199; ZF_RANBP2_2; 1

PROSITE; PS50199; ZF_RANBP2_2; 1
 PubMed=3031652;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL; U13178; AAA86955.1; -.
EMBL; L37083; AAC41563.1; -.
EMBL; AE003501; AAE48578.1; -.
EMBL; M15765; AAA70425.1; -.
FlyBase; FBgn0011571; caz.
InterPro; IPR000504; RRM.
InterPro; IPR001876; Znf-RanBP.
Pfam; PF00076; rrm; 1.
Pfam; PF00641; zf-RanBP; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    41
92
108
258
M
283
283
D
398
D
39141 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           23.0%; 43.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          97 555555--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         KVSLAQRQNNWNKGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       205
205
275
309
391
47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      splicing
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
MEDLINE=87175568;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Nuclear protein;
Alternative splic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 42
1119
2122
2122
280
39
4
39
39
108
2253
389
389
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               -- 555555
                                                                                                                                                                                                                                                                                                                                                                                                             DETECTED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ZN_FING
DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         199
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                    THE WARRANG BRANK 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ò
```

4;

238 GGGGGRFDRGGGGGG

g

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                  ostomi;
te; Mus.
                                                                                                                                                                                                      ey D.M
                                                                                                                                                                                                                              a new
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7.
                                                                                (Fragment).
                                                                                                                   Euteleo:
                                                                                                                                                                     SEQUENCE FROM N.A.
STRAIN=BALB/C; TISSUE=Liver;
MEDLINE=94195427; PubMed=8145850;
Storm E.E., Huynh T.V., Copeland N.G., Jenkins N.A., Kingsl
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         GROWTH/DIFFERENTIATION FACTOR BY SIMILARITY.
BY SIMILARITY.
BY SIMILARITY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 151
                                                                                                                                                                                                                               in
                                                                                                                                                                                                                   Lee S.-J.;
"Limb alterations in brachypodism mice due to mutations i member of the TGF beta-superfamily.";
Nature 368:639-643(1994).
-! SUBUNIT: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
-! SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            BY SIMILARITY.
INTERCHAIN (BY SIMILARITY)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    0E496AACB5827759 CRC64;
                                                                                    Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; E
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
NCBI_TaxID=10090;
                     GDF7_MOUSE STANDARD; PRT; 151 AA. P43029; 01-NOV-1995 (Rel. 32, Created) 01-NOV-1995 (Rel. 32, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Growth/differentiation factor 7 precursor (GDF-7)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 125; DB 1; Pred. No. 6.7e-05; Hismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                EMBL; U08339; AAA18780.1; --
HSSP; P12643; 3BMP.
MGD; MGI:95690; Gdf7.
InterPro; IPR001839; TGF-beta.
ProDom; PD000357; TGF-beta; 1.
PROSITE; PS00250; TGF_BETA_1; 1.
Growth factor; Cytokine; Glycoprotein.
NON_TER 1 1 1
PROPEP <1 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2002, 09:28:20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       POLY-ARG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     POLY-G
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         151
116
148
150
115
5 PP
41
1597 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        22.9%;
54.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     October 24,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 Similarity
25; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                <1
<6
50
79
83
1115
116</pre>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 151
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Search completed: C
Job time: 14.6788
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DISULFID
DISULFID
DISULFID
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DISULFID
RESULT 15
GDF7_MOUSE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          34
                                                           ŏ
```

secs

				•	

.rspt

us-09-894-030-3

```
updates/sec
                                                                                                                                                                                                          Seconds
                                                                                                                                                                                                                                                                                                                                            US-09-894-030-3
545
1 MEKLFKEVKLEELENQKGSG.....DDFDLDVVKVSKQDSKITPQ 100
                                                                                                                                                                                                      ; Search time 32.1168 (without alignments) 538.643 Million cell
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            562222
5.1.3
Compugen Ltd
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Total number of hits satisfying chosen parameters:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           562222 seqs, 172994929 residues
   GenCore version Copyright (c) 1993 - 2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                                                           using sw model
                                                                                                                                                                                                            2002, 09:24:21
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          sp_archea:*
sp_archea:*
sp_bacteria:*
sp_human:*
sp_invertebrate:*
sp_mammal:*
sp_mhc:*
sp_phage:*
sp_phage:*
sp_phage:*
sp_rodent:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_virus:*
sp_vartebrate:*
sp_virus:*
sp_vartebrate:*
sp_vartebrates:*
sp_var
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      0.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     BLOSUM62
Gapop 10.0 , Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   seq length: 0 seq length: 2000000000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SPTREMBL_19:*
                                                                                                                                           - protein search,
                                                                                                                                                                                                            October 24,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Title:
Perfect score:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Scoring table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Minimum DB
Maximum DB
                                                                                                                                             OM protein
                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Searched:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Database
                                                                                                                                                                                                              Run on:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Description	4033	O34781 bacillus su	Q93424 caenorhabdi	Q42992 oryza sativ	Q9bzg5 homo sapien	homo	Q9bzg6 homo sapien	homo		Q38777 allium sati	Q92nu7 rhizobium m	Q18880 caenorhabdi	O65330 elaeagnus u	Q43522 lycopersico	Q9grw7 drosophila	Q91trl tupaia herp
DB ID	9 064033	16 034781	5 093424	10 042992	4 Q9BZG5	4 Q9NUA2	4 Q9BZG6	4 Q9BZG7	10 065514	10 038777	16 Q92NU7	5 Q18880	10 065330	10 043522	5 Q9GRW7	12 Q91TR1
Length D	26	26	385	333	531	539	542	544	221	318	233	393	335	207	697	718
% Query Match			27.7	27.0	26.1					25.9	٠	•	•	•	25.1	25.0
Score	. 0	304		147	42.	142.5	42.	142.5	4	141	9.	139.5	C	S	137	136.5
Ω T		2	e	4	Ŋ	9	7	8	6			12				16

25 24.8 163 5 Q95UX4 Q95ux4 drosophila Q95ux4 drosophila Q95nK6 165 5 Q95NK6 Q95NK6 Q95NK6 Grosophila Q95nK6 Grosophila Q95nK 165 5 Q95NK2 Q95NK6 Grosophila Q95NK2 Grosophila Q95NK2 Q95NK2 Grosophila Q95NK2 Q95NK2 Q95NK2 Grosophila Q95NK2 Q93010 Oryza sativ Q95NK1 11 Q99NK1 Q99NK1 Q99NK1 Q99NK1 Q99NK1 G043Q4 Oryza sativ Q95NK2 1163 5 Q95NW3 Q95NW9 Q95NW9 Grosophila Q95NW9 Grosophila Q95NW9 Grosophila Q95NW9 Grosophila Q95NW9 Grosophila Q95NW Grosophila Q95NW9 Grosophila Q95NW	PRELIMINARY; PRT; 56 AA.  1998 (TrEMBLrel. 07, Created) -1998 (TrEMBLrel. 07, Last sequence update) -2001 (TrEMBLrel. 19, Last annotation update) -2001 (TrEMBLR	E FROM N.A. ic V., Duesterhoeft A., Soldo B., Hi a D.; mplete nucleotide sequence of the Ba e."; ed (AUG-1997) to the EMBL/GenBank/DD F020713; AAC12992.1; tein. E 56 AA; 5982 MW; 79ECOBF822F9F4 h 55.8%; Score 304; DB Similarity 100.0%; Pred. No. 5.8 56; Conservative 0; Mismatches	THE FRENCH ELEMENTS OF THE PRODUCT OF THE PRODUCT OF THE PRESENCE OF THE PRESE
17 135 19 135 19 135 20 134.5 22 134.5 22 134.5 24 134.5 25 133.5 26 130.5 31 130.5 33 130.5 34 130.5 35 130.5 36 130.5 37 130.5 38 129.5 42 129	T 1 3 064033 064033; 01-AUG-1 01-AUG-1 01-DEC-2 PUTATIVE YOLG. Bacterio Viruses;	NCB1_Tax. [1] SEQUENCE Lazarevic Karamata The comprophage Submitted EMBL; AF Lipoprott SEQUENCE SEQUENCE St Local tches 5	LT 2 81 034781 034781 01-JAN-1 01-JAN-1 01-JAN-1
	RESUL 06403 ID AC DT DT DT OS OS	CX RP RRA RA RT RT RW SQ Ouc	Db RESU 0347 ID AC DT DT

Thu

.rspt

```
[1]
SEQUENCE
                                                                                                                                                                                                 REVISIONS
                                                                                                                                                                                                                                                                                                                                                                       TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             141
                                                                                                                                                                                7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
042992
           ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ö
                                                                                                                                                                                                 SECURE FROM N.A.

STRAINIES,
WEDLINE-98044033; PubMed-9384377;
WEDLINES R., Boursier L., Brans A., Braun M., Brignell S.C., Bron. S., Chois S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoff A., Ehrlich S.D., Emmerson P.T., Rabritz C., Fujita M., Fabret C., Ferrari E., Foulger D., Rohis S., Galiszia A., Galleron N., Chim S.Y., Glaser P., Goffend A., Golightly E.J., Grandi G., Rightly E.J., Garamata D., Kashara Y., Klaerr-Blanchard M., Klein C., Mosysshi Y., Koetter P., Koningstein G., Krogh S., Kumano M., Joris B., Karamata D., Kashara Y., Klaerr-Blanchard M., Klein C., Mosysshi Y., Koetter P., Koningstein G., Krogh S., Kumano M., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maudel C., Medigue C., Medigue C., Medigue C., Oth T.M., Porteclel D., Porwollik S., Perescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Sator V., Pohl T.M., Porteclel D., Porwollik S., Perescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Tamakoshi A., Tamaka H., Takahashi H., Takamaru K., Takahashi H., Takamaru K., Takahashi H., Takamaru K., Takamaru K., Takahashi H., Takamaru K., Takahashi H., Takamaru K., Mandut R., Wander E., Wedler E., Wedler E., Wedler E., Wedler E., Weller E., Wollida K., Yoshikawa H.F., Zumstein E., Yoshikawa H.F., Panchin A., Tamakoshi A., Tamaka H., Yoshikawa H.F., Zumstein E., Yoshikawa H.F., Zumstein E., Yoshikawa H.F., Zumstein E., Politius S., Fall F., Sall F.,
                                                                                                                                                                                                              Alloni G.,
Borchert S.,
ell S.C., Bron S.,
'., Carter N.M.,
Daniel R.A.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 in A.;
Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       \alpha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SOME
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                56
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   IN CHROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                         STRAIN=168;
Paik S.H., Hansen J.N.;
Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; AF014938; AAC63531.1; -.
EMBL; 299115; CAB14066.1; -.
Complete proteome.
SEQUENCE 56 AA; 5982 MW; 79EC0BF822F9F4C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   79EC0BF822F9F4C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              093424 PRELIMINARY; PRT; 385 AA. 093424; P90801; 01-JAN-1999 (TrEMBLrel. 09, Created) 01-JAN-1999 (TrEMBLrel. 09, Last sequence update) 01-DEC-2001 (TrEMBLrel. 19, Last annotation update) HYPOTHETICAL GLYCINE-RICH 37.0 KDA PROTEIN E02A10.2
                                 Bacillus subtilis.
Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 304; DB 16;
Pred. No. 5.8e-26;
0; Mismatches 0;
SUBLANCIN 168 PRECURSOR PEPTIDE SUNA.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                55.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5982 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Nature 390:249-256(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
56; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FROM N.A.
                                                                                                               FROM N.A
                                                                                                                                                                                         SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STRAIN=168;
                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT
Q93424
                                                                                                                                                                                       E H H A C P
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            à
```

```
Rhabditida; Rhabditoidea;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CHITINASE PRECURSOR (EC 3.2.1.14).
Oryza sativa (Rice).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

STRAIN=IR36; TISSUE=ETIOLATED LEAF;

A Yun C.-H., Kim J.-K., Park Y.-H.;

T "Isolation and Characterization of A Rice Chitinase cDNA.";

L Submitted (NOV-1994) to the EMBL/GenBank/DDBJ databases.

R HSSP; P23951; ZBAA.

R HSSP; P23951; ZBAA.

InterPro; IPR001002; Chitin_bind.

R HSSP; P23951; Chitin_binding; 1.

R Ffam; PF00187; chitin_binding; 1.

PRODM; PR00451; CHITINBINDNG.

DR PRODM; PD000674; Glyco_hydro_19; 1.

PRODM; PD000609; Chitin_bind; 1.

DR SMART; SM00270; ChtBDI; 1.

DR PROSITE; PS00773; CHITINASE_19_1; 1.

DR PROSITE; PS00774; CHITINASE_19_2; 1.

DR PROSITE; PS00026; CHITINASE_19_2; 1.

PROSITE; PS00026; CHITINASE_19_2; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             GLYCINE-RICH PROTEIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 385;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                        databases
                                                                                                                                                                                                                                                                                                                                              databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CHITINASE.
E2B65D4155B8C73D CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       33 CASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              E02A10.2.
POTENTIAL.
GLY-RICH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18;
16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 151; DB 5;
Pred. No. 1.8e-08
; Mismatches 10
                                                                                                                                                                                                                       to the EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                                                                              to the EMBL/GenBank/DDBJ
                                                                                                                                                                                                                                                                                                                                                                                                                                              Transmembrane
                                              da; Chromadorea;
Caenorhabditis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  POTENTIAL.
HYPOTHETICAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                333
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                       Submitted (JUN-1997) to the EMB EMBL; Z81053; CAB02877.1; -. HSSP; P10968; 2CWG. WormPep; E02A10.2; CE09116. Hypothetical protein; Signal; TSIGNAL 18 385 HY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       185 P
190 G
36963 MW;
                                        Eukaryota; Metazoa; Nematoda;
Rhabditidae; Peloderinae; Caer
NCBI_TaxID=6239;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         27.7%;
63.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRELIMINARY;
E02A10.2.
Caenorhabditis elegans
                                                                                                                                                                                       Thomas K.;
Submitted (OCT-1996)
                                                                                                                                           SEQUENCE FROM N.A. STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Matcn
Best Local Similarity
                                                                                                                                                                                                                                                                               REVISIONS.
STRAIN-BRISTOL N2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        385 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         NCBI_TaxID=4530;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         165
23
                                                                                                                                                                                                                                                                                                                   Jones S.J.M.;
Submitted (JU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     042992;
042992;
01-NOV-1996 (
```

ö

g

δ

g

'n

```
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     476
                                                                                                                                                                                                                                        473
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GGG 76
                                                                                                                                                                                                                  . 74 GGG 76
                                                                                                                                           32;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ---
666
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
Best Local (
                                                                                                 SEQUENCE
                                                                                                                                                                                                                                         GGG
                                                                                                                        Query Match
                                                                                                                                   Local
                                                                                      NON_TER
                                                                                                                                                                                                                                                                                                 Q9BZG6
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Q9BZG7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best Loc
Matches
                                                                                                                                   Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    474
                                                                                                                                                                      18
                                                                                                                                                                                                                                        471
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                                       99ZE60
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Q9BZG7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                RL
DR
DR
DR
DR
FT
                                                                                                                                                                                            g
                                                                                                                                                                                                                                         a
                                                                                                                                                                                                                                                                                                   П
                                                                                                                                                                      ò
                                                                                                                                                                                                                  ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                              4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        462
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                     --- CGGGAVACQNYRQFCRGGGGGGGGGGGGGG 70
                                                                 11111111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        e breast
                                                                                                                                                                                                                                  ANDROGEN RECEPTOR (FRAGMENT).
Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     eostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            16;
                              12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                      531;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Q9NUA2
Q9NUA2;
Q9NUA2;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DJ80804.1 (ANDROGEN RECEPTOR (DIHYDROTESTOSTERONE RECEPTOR))
        333;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Ŧ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               111111111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                              Indels
        Length
                                                                                                                                                                                                                                                                                                                     Magklara A., Diamandis E.P.;
"Poly-Q and poly-G repeats in the androgen receptor of tancer cell line MCF-7.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases EMBL; AF321917; AAK09426.1; -.
InterPro; IPR001103; Androgen_recep.
Pfam; PF02166; Androgen_recep; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      of
                                                                                                                                                                                                                                                                                                                                                                                                                              FA0D3AA60AC20DAB CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.
Magklara A., Diamandis E.P.;
"Poly-Q and poly-G repeats in the androgen receptor
                                                                                                                                                                                                     (TrEMBLrel. 17, Created)
(TrEMBLrel. 17, Last sequence update)
(TrEMBLrel. 19, Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                             16;
      10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 142.5; DB Pred. No. 2.2e-07
      Score 147; DB 10
Pred. No. 4.3e-08;
; Mismatches 10
                                                                                                                                                                                531 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               GSGSPSAAASSSWHTLFTAEEGQLYGPCGGG--
                                                                                                                                                                                PRT;
                                                     16 QKGSGLGKAQCAALWLQCASGGTIG--
                            2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            .,
m
                                                                                                                                                                                                                                                                                                                                                                                                                             54419 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                   26.1%;
illarity 50.8%;
Conservative
      27.0%;
larity 54.5%;
Conservative
                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.
Chapman J.;
Submitted (APR-2000)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (Human)
Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                             531 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local Similarity
nes 32; Conser
                                                                                                                                                                                                                                                                                                               N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           NCB1_TaxID=9606;
                                                                                                 71 GGGGGG 76
                                                                                                                                                                                                                                                                                                             FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     463 GGG 465
                                                                                                                                                                                                     01-JUN-2001
01-JUN-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              16
                                                                                                                        78 GGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (FRAGMENT)
                                                                                                                                                                                                                                                                                                                                                                                                         Receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                   NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                999
                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                Q9BZG5
Q9BZG5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       415
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                                                                                                                                                           RESULT
                                                                                                                                                                      Q9BZG
```

RA RT RL DR

DR DR KW FT SO

g

ö

g

οy

```
470
                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18 GSGLGKAQCAALW--LQCASGGTI--GCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        of the breast
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Euteleostom1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   16;
                                                                                                                                                                                                                                         16;
                                                                                                                                                                                                   539;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
cancer cell line T-47D.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases
EMBL; AL049564; CAB87955.1; -.
EMBL; AF321915; AAK09424.1; -.
InterPro; IPR001103; Androgen_recep.
Pfam; PF02166; Androgen_recep; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   C363EF841CAF7739 CRC64;
                                                                                                                                                            AB493953B89D869F CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
Magklara A., Diamandis E.P.;
"Poly-Q and poly-G repeats in the androgen receptor cancer cell line BT-474.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databa EMBL; AF321916; AAK09425.1; -.
InterPro; IPR001103; Androgen_recep.
Pfam; PF02166; Androgen_recep; 1.
Receptor.
NON_TER 542 542
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
ANDROGEN RECEPTOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               01-JUN-2001 (TrEMBLrel. 17, Created)
01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
ANDROGEN RECEPTOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 12;
                                                                                                                                                                                                 Score 142.5; DB 4;
Pred. No. 2.2e-07;
3; Mismatches 12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 142.5; DB 4 Pred. No. 2.2e-07;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           542 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                         423 GSGSPSAAASSSWHTLFTAEEGQLYGPCGGG---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               426 GSGSPSAAASSSWHTLFTAEEGQLYGPCGGG--
                                                                                                                                                                                                                                        3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 3;
                                                                                                                                         539
55444 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 542
55828 MW;
                                                                                                                                                                                                 26.1%;
illarity 50.8%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26.1%;
50.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                542
542 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 Similarity
32; Conserv
                                                                                                                                         539
539 AA;
                                                                                                                                                                                                                         Similarity
```

3

D.

```
PEQUENCE FROM N.A.

SEQUENCE FROM N.A.

Van Damme E.J.M., Willems P., Peumans W.;

Van Damme E.J.M., Willems P., Peumans W.;

Yao Damme E.J.M., Willems P., Peumans W.;

Tisolation and characterization of two different chitinase cDNA clones

If rom garlic (Allium sativum L.) shoots.";

Submitted (MAY-1992) to the EMBL/GenBank/DDBJ databases.

REMBL, M94105; AAA32641.1; -.

REMBL, M94105; AAA32641.1; -.

REMBL, M94105; AAA32641.1; -.

REMBL, M94105; Chitin_bind.

RICEPTO; IPR001002; Chitin_bind.

REMBL, M960187; Chitin_bind; 1.

REMBL, M960182; Glyco_hydro_19; 1.

REMBL, M960182; Glyco_hydro_19; 1.

REMBL, M960182; CHITINASE_19; 1.

REMBL, M960183; CHITINASE_19-1; 1.

REMBL, M960193; CHITINASE_19-1; 1.

REMBL, M960194; CHITIN_BINDING; 1.

REMBL, M960195; CHITIN_BINDING; 1.

REMBL, M960195; CHITIN_BINDING; 1.
                                                        Allium sativum (Garlic).
Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;
  01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
CHITINASE PRECURSOR (FRAGMENT).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 LFKEVKLEELENQKGSGLGKAQCAALWLQCASGGTIG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            GGGGGGGGGGGGGGGMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local Similarity
Les 37; Conserv
                                                                                                                      NCBI_TaxID=4682;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11 LFK--
                                              CHITINASE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                092NU7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           59
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
Q92NU7
    g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Tracheophyta;
dicots; Rosidae;
                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                     47
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft
Hoheisel J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.,
Schueller C.;
                                                                                                                                                                                                                                                                                   --6666666666666666
                                                breast
                                                                                                                                                                                                                           15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ö
                                                                                                                                                                                              544;
                                                                                                                                                                                                                                                                                                                                                                                                                    ob5514;
01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-AUG-1998 (TrEMBLrel. 15, Last annotation update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
PUTATIVE GLYCINE-RICH CELL WALL PROTEIN.
F23E13.120 OR AT4G36230.
Arabidopsis thaliana (Mouse-ear cress).
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheol Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             221;
                                                                                                                                                                                               Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft
Mewes H.W., Lemcke K., Mayer K.F.X.;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                          Magklara A., Diamandis E.P.;
"Poly-Q and poly-G repeats in the androgen receptor of t cancer cell line 2R-75-1.";
Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AF321914; AAK09423.1; -.
InterPro; IPR001103; Androgen_recep.
Pfam; PF02166; Androgen_recep.
Receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      EU Arabidopsis sequencing project;
Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases
EMBL; AL022141; CAA18129.1; -.
EMBL; AL161589; CAB80294.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Submitted (MAR-1998) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                 C97133EB6C922E74 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Score 141; DB 10; I
Pred. No. 1.2e-07;
                                                                                                                                                                                             4;
                                                                                                                                                                                            Score 142.5; DB 4;
Pred. No. 2.2e-07;
3; Mismatches 13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DE8DAE2E2C57A9F8
                                                                                                                                                                                                                                                                                                                                                                                                                  221 AA
                                                                                                                                                                                                                                                                       PRT;
                                                                                                                                                                                                                          э;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                              56013 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20344 MW;
                                                                                                                                                                                            26.1%;
larity 50.8%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25.9%;
llarity 65.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                 PRELIMINARY;
                                                                                                                                                              544 AA;
                                                                                                                                                                                                           Similarity 32; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         N.A.
                N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Similarity
28; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                EQUENCE FROM N.A.
 [1]
SEQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM
EU Arabidopsis
                                                                                                                                                                                                                                                                                                                                         GGG 478
                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Mewes H.W.,
Submitted (1
                                                                                                                                                                                                                                                                                                              999
                                                                                                                                                               SEQUENCE
                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE
                                                                                                                                                                                                          Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       34
                                                                                                                                                                                                                                                                                                                                         476
                                                                                                                                                                                                                                                                                 427
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 130
                                                                                                                                                                                                                          Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                   RESULT
065514
                                                                                                                                                                                                            Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Best
REART REARING DON DON KW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        qq
                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                      ò
                                                                                                                                                                                                                                                                                                              ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  a
```

ŭ,

Length 318;

POTENTIAL.
CHITINASE.
097C3DBD772468D4 CRC64;

33175 MW;

318 AA;

318

Indels

Score 141; DB 10; Pred. No. 1.8e-07; 3; Mismatches 22

5,

Conservative

25.9%; 46.2%;

--CGGGAVACQNYRQFCRGG 58

--NSYAQQCGSQAGGALCSNR-LCCSKFGYCGSTDPYCGTG---CQSQ---CGGG

```
Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F., Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G., Boutry M., Bowser L., Buhrmester J., Cadieu E., Capela D., Chain P., Cowie A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F., Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M., Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D., Kahn M.L., Kalman S., Keating D.H., Kiss E., Komp C., Lelaure V., Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B., Ramsperger U., Surzycki R., Thebault P., Vandenbol M.,
                                                                                                                                                                                                                                                               Rhizobium meliloti (Sinorhizobium meliloti).
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Rhizobiaceae; Sinorhizobium.
                                                                                                                                      092NU7;
01-DEC-2001 (TrEMBLrel. 19, Created)
01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
PUTATIVE GLYCINE-RICH PROTEIN.
                                                                                                                         233 AA.
                                                                                                                                                                                                                                                                                                                                                                                                          MEDLINE=21368234; PubMed=11474104;
PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                   FROM N.A.
                                                                                                                                                                                                                                                                                                        Rhizobiaceae; S
NCBI_TaxID=382;
                                                                                                                                                                                                                                                                                                                                                                                        STRAIN=1021
                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE
```

**S** 

Created)

01,

(TrEMBLrel.

038777; 01-NOV-1996

038777

RESULT 10 Q38777

Thu

KW SO

DR

8

ð

ı

4

```
GGGG 74
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       111 |
GSGSG--
                                                                                                                                                                Query Match
Best Local Sim
                                                                                                                                                                                                                                                                                ---
2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                  71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              85
                                                                                                                                                                                                                   16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
Q9GRW7
                                                                                                                                                                                                                                                                                                                              RESULT
Q43522
                                                                                                                                                                                                                                                                                                                                                                                       ð
g
                                                                                                                                                                                                                                                                   ŏ
                                                                                                                                                                                                                                                                                          g
                                                                                                                                                                                                                     õ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5
                                                                                                          2;
138
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Tracheophyta;
Idicots; Rosidae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Gaps
                                                                                                            Gaps
                                                                                                                                 76
                                                                                                                                                                                                                                                                                                                   bditoidea;
  ., Batut J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16
                                                                                                                                               1111111111111111111
                                                                                                                                   -GGCCGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18 GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  13;
                                                                                                          19;
                                                                                   233;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             dic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          393;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ts;
S., Wells D.H., Wong K., Yeh K.-C the legume symbiont Sinorhizobium
                                                                                                                                                                                                                                                                                                                                                                                                                                                    for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SECUENCE FROM N.A.
TISSUE=ROOT NODULE;
Kim H.-B., An C.-S.;
"Isolation and characterization of cDNA clone encoding aci
chitinase from the root nodule of Elaeagnus umbellata.";
Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
EMBL; AF061805; AAC16010.1;
                                                                                                                                                                                                                                                                                                                    Nematoda; Chromadorea; Rhabditida; Rha
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracl
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot
eurosids I; Rosales; Elaeagnaceae; Elaeagnus.
NCBI_TaxID=43233;
                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Indels
                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                         Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                   A platform
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2843994C2F6DE0F0 CRC64;
                                                           ED22171714E58E55 CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          065330;
065330;
01-AUG-1998 (TrEMBLrel. 07, Created)
01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
ACIDIC CHITINASE (EC 3.2.1.14).
                                                                                                                                                                                                                                                        Last sequence update)
Last annotation update)
                                                                                 DB 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 139.5; DB 5;
Pred. No. 3.3e-07;
0; Mismatches 16;
                                                                                                 )7;
10;
                                                                                  Score 139.5; DB Pred. No. 1.9e-07, Mismatches 10
                                                                                                                                                                                                                                                                                                                                                                                                                                              "Genome sequence of the nematode C.elegans:
investigating biology.";
Science 282:2012-2018(1998).
EMBL; Z77655; CAB01137.1; ..
SEQUENCE 393 AA; 37490 MW; 2843994C2F6DE
                                                                                                                                                                                                                                                                                             C56A3.1.
Caenorhabditis elegans.
Eukaryota; Metazoa; Nematoda; Chromadorea
Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
                                                                                                                                                                                                                     393
                                                                                                                                                           98 RGGGKGGGS-----KGGGGVGGGGGG----
                                                                                                                                                                                                                                             Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        87 GGGGGGGGGGGGGGGG--GCGGGG-
                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.
MEDLINE-99069613; Pubmed-9851916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ô
                                                                                                         2;
                                                          22513 MW;
Vorhoelter F.J., Weidner S.,
"The composite genome of the
Science 293:668-672(2001).
EMBL; AL591789; CAC46652.1; -
Complete proteome.
SEQUENCE 233 AA; 22513 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25.6%;
50.8%;
                                                                                   25.6%;
48.3%;
                                                                                                                                                                                                                                           01,
05,
19,
                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Conservative
                                                                                                                                                                                                                                           (TrEMBLrel. (TrEMBLrel. (TrEMBLrel.
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity
                                                                                              Local Similarity
nes 29; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE FROM N.A TISSUE=ROOT NODUL
                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                C56A3.1 PROTEIN.
                                                                                                                                                                                                                                           01-NOV-1996
01-JAN-1998
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                30;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
Best Local
                                                                                   Query Match
                                                                                                                                                                                                                    Q18880
Q18880;
                                                                                                                                                                                                                                                                                                                                                                              Sims
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 13
                                                                                                                                                                                               RESULT
018880
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    121
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              92
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       enhanced
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Tracheophyta;
                                                                                                                                                                                                                                                                                                                                                                          OKGSGLGKAQCAALWLQCASGGTIG-----CGGGAVACQNYROFCRGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              GSGLGKAQCAALWLQCASGGTIGCGGGAVACQNYRQFCRGGGGGGGGGGGGGGGGGGGGGGGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 22;
                                                                                                                                                                                                                                                                                                                  12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Lycopersicon esculentum (Tomato).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Trachec
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
NCBI_TaxID=4081;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      two tomato fruit
HSSP; P23951; 2BAA.
InterPro; IPR001002; Chittin_bind.
InterPro; IPR000726; Glyco_hydro_19.
Pfam; PF00187; Chittin_binding; 1.
Pfam; PF00182; Glyco_hydro_19; 1.
PRINTS; PR00451; CHITINBINDNG.
ProDom; PD000574; Glyco_hydro_19; 1.
SMART; SM00270; ChtBDl; 1.
PROSITE; PS00774; CHITINASE_19_1; 1.
PROSITE; PS00076; CHITINASE_19_2; 1.
PROSITE; PS00026; CHITIN_BINDING; 1.
Chitin-binding; Glycosidase; Hydrolase.
SEQUENCE 335 AA; 35737 MW; 1CC09BB8767A18E0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A50FE71F8ED4C4D9 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score 138; DB 10;
Pred. No. 2.4e-07;
Mismatches 6;
                                                                                                                                                                                                                                                                                                                16;
                                                                                                                                                                                                                                                                          Score 139; DB 10;
Pred. No. 3.2e-07;
                                                                                                                                                                                                                                                                                               ed. No. 3.26
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN-CV. UC82B; TISSUE-FRUIT;
MEDLINE-97201476; PubMed-9049262;
Santino C.G., Stanford G.L., Conner T.W.;
"Developmental and transgenic analysis of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      -GGGSGTGGGGGS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A.
STRAIN=CV. UC82B; TISSUE=FRUIT;
Connor T.W.;
Submitted (JAN-1996) to the EMBL/
EMBL; X95262; CAA64559.1; -.
HSSP; P24337; 1HYP.
InterPro; IPR003612; AAI.
InterPro; IPR000528; Plant_LTP.
Pfam; PF00279; LTP; 1.
SMART; SM00499; AAI; 1.
SEQUENCE 207 AA; 18265 MW; A<sup>±</sup>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           genes.";
plant Mol. Biol. 33:405-416(1997)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                   3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25.38;
50.8%;
                                                                                                                                                                                                                                                                         n
Similarity 51.6%;
33; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             01,
01,
19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q43522
Q43523;
01-NOV-1996 (TrEMBLrel. 0
01-NOV-1996 (TrEMBLrel. 0
01-DEC-2001 (TrEMBLrel. 1
TFM5 PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 50.8 30; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PRELIMINARY;
```

.; ;;

덛

g

à

RR RP RC

```
Figure From N.A.

Campesan S., Chalmers D., Sandrelli F., Megighian A., Peixoto A.A.,
Campesan S., Chalmers D., Sandrelli F., Megighian A., Peixoto A.A.,
Costa R., Kyriacou C.P.;
Comparative analysis of the nonA region in Drosophila identifies
I "Comparative analysis of the nonA region in Drosophila identifies
I highly diverged 5' gene that may constrain nonA promoter evolution.";
L submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
EMBL; AJ298998; CAC10058.1; -.
R HSSP; P11940; ICVJ.
R HSSP; P11940; ICVJ.
R FlyBase; FBgn0042729; Dvir\nonA.
InterPro; IPR000504; RRM.
R FlyBase; FBGn00450504; RRM.
SMART; SM00360; RRM; 2.
R FROSITE; PS50102; RRM; 2.
R PROSITE; PS50102; RRM; 2.
R PROSITE; PS50102; RRM; 2.
                                                                                                                    Drosophila virilis (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     12; Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 697;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
Q9GRW7;
Q9GRW7;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
NONA PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 137; DB 5; Pred. No. 1.1e-06; 1; Mismatches 8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
Best Local Similarity 57.1%;
Matches 28; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          NONA
    qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
```

Search completed: October 24, 2002, 09:29:12 Job time : 35.1168 secs

```
5.1.3
Compugen Ltd
 version - 2002
GenCore
Copyright (c) 1993
```

using sw model protein search, OM protein

Seconds 09:23:41 2002, 24, October Run on:

updates/sec ; Search time 14.3139 (without alignments) 287.115 Million cell

Title:

MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG US-09-894-030-3\_COPY\_1\_37 193 1 MEKLFKEVKLEELENOKGGGIGK score: Sequence: Perfect

37

BLOSUM62 Gapop 10.0 , Scoring table:

111073796 residues 747574 seqs, Searched:

Gapext 0.5

747574 Total number of hits satisfying chosen parameters

2000000000 seq length: seq length: Minimum DB Maximum DB

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database

19. DAT: \*
10. DAT: \*
11. DAT: \*
12. DAT: \*
13. DAT: \*
14. DAT: \*
15. DAT: \*
16. DAT: \*
17. DAT: \*
18. DAT: \*
19. DAT: \* DAT: \*
DAT: \*
DAT: \*
DAT: \*
DAT: \*
DAT: \*
DAT: \*
DAT: \*
DAT: \* A\_ceneseq\_032602:\*

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT

SIDS1/gcgdata/geneseq/geneseqp-emb1/AA1999 A\_Geneseq\_032802 

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

	Description		Amino acid sequenc	Amino acid sequenc	Amino acid sequenc	Arabidopsis thalia	Arabidopsis thalia	Cystathionine beta	Arabidopsis cystat	Arabidopsis thalia	Arabidopsis thalia	Arabidopsis thalia	Arabidopsis thalia
SUMMARIES	ID		AAY03205	AAY03207	AAY03218	AAG16276	AAG16275	AAB47604	AAB20017	AAG16274	AAG58269	AAG60923	AAG58268
	DB	;	20	20	20	21	21	22	22	21	21	21	21
	Length		56	37	37	408	464	464	464	506	200	200	202
	& Ouery Match	                 	100.0		49.7	30.1	30.1	30.1	30.1	30.1	29.0	29.0	29.0
	Score		193		96	58	58	28	28	58	26	56	26

Result No.

•			(	ć	(	:
12		J.	$\circ$	7.7	609	s thal
13		9	-	21	092	dopsis thal
14		δ		21	AAG58267	Arabidopsis thalia
15	Ŋ.	8		20	12	de
16			84	22	3	Cone snail O-supe
17	S	œ	0	21		Human secreted pro
18	S		4	22	2814	
19	4	ω	3	20	015	A human vanilloid
20	54.5	28.2	839	20	30	ial huma
21	4	8	$\sim$	20	015	l vanil
22	4.	8	æ	20	0655	i.C
23	4.	8	3	21	3212	$\overline{}$
24	4.	æ	$\mathcal{C}$	21	647	nilloid
25	4.	8	3	21	735	Д
26	4	8	3	22	122	n vanil
27	ω.	1		22	593	
28	ж	7	0	22	3	la melan
29	Ŋ		6	18	900	a t
30	S	7	9	20	102	ent of
31		6	0	22	939	
32	Ŋ	9	7	20	101	S
33	1	9	æ	22	496	24 human s
34	1.	9	378	22	96	Gene 24 human secr
35	۲,	9	4	22	813	human d
36	。	9	63	18	038	lori deri
37	。	9	9	22		
38			786	22		Human polypeptide
39	0	9	4	22	561	Human protein sequ
40	0	9	σ	22	~	Human polypeptide
41		Ŋ		22	~	Novel human secret
42		2	7	22	900	Human polypeptide
43		Ŋ	375	20	ø	ano
44		ហ		22	78	Human polypeptide
45		2		21	$\sim$	Human secreted pro
					ALIGNMENTS	
RESULT 1						
ID AAYO	05	standard;	Protein;	iu;	56 AA.	

AAY03205; 

(first entry) 03-AUG-1999

sunA protein. of sednence Amino acid

pro-sublancin 168; sunT. Sublancin 168; antimicrobial; food preservative; Gram-positive bacteria; pre-sublancin 168; sunA;

Bacillus subtilis

WO9903352-A1

28-JAN-1999

17-JUL-1998;

97US-0053035 98WO-US14547 18-JUL-1997; (UYMA-) UNIV MARYLAND BALTIMORE

Hansen JN;

WPI; 1999-131752/11 N-PSDB; AAX28631.

New antimicrobial peptide, sublancin 168, from Bacillus subtilis used for, e.g. treatment of infections caused by Gram negative bacteria and as food preservative

a

ŏ

sublancin 168

of

(first entry)

AA

37

37

standard; Protein;

```
Sublancin 168; antimicrobial; food preservative; pro-sublancin 168; Gram-positive bacteria; Pro-sublancin 168; pre-sublancin 168.
GLGKAQCAALWLQCASGG
               GLGKAQCAALWLQCASGG
                                                                                                                                                                                  Amino acid sequence
                                                                                                                                                                                                                                                        Bacillus subtilis
                                                                                                                                                       03-AUG-1999
                                                                                                                           AAY03218;
                                                                                               AAY03218
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WPI;
20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                    RESULT 3
                                                                                                                                                                                                                                                                                     Key
                                                                                                                                                                                                                                                                                                                                                      Qγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        d
ŏ
                                                                                                                                                                                                                            ö
                                                                                                                                                                                                                           Gaps
                        This is the amino acid sequence of the sunA protein encoded by the prosublancin 168 nucleotide sequence used in the method of the invention. The peptide designated sublancin 168, is an antimicrobial useful for treating infections and preserving food against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years in aqueous solution of about
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This is the amino acid sequence of prosublancin 168 used in the method of the invention. The peptide designated sublancin 168, is an antimicrobial useful for treating infections and preserving food against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years in aqueous solution of about neutral pH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Sublancin 168; antimicrobial; food preservative; pro-sublancin 168; Gram-positive bacteria; pre-sublancin 168.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         sublancin 168, from Bacillus subtilis
                                                                                                                                                                                                                           ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     antimicrobial peptide, sublancin 168, from Bacillus subt
1 for, e.g. treatment of infections caused by Gram negati
eria and as food preservative
                                                                                                                                                                                                56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 37;
                                                                                                                                                                                                 Length
                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 20;
3.1e-06;
0;
                                                                                                                                                                                              ; DB 20;
1.9e-18;
hes 0;
                                                                                                                                                                                                                                                        37
                                                                                                                                                                                                                                                                                 37
                                                                                                                                                                                                                                                       1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                                                                                                                                                                                                    red. No. 1. Mismatches
                                                                                                                                                                                               Score 193;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 100;
                                                                                                                                                                                                                                                                                                                                                                                                                                         of prosublancin 168
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Claim 1; Pages 52-53; 71pp; English.
                                                                                                                                                                                                               Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MARYLAND BALTIMORE
                                                                                                                                                                                                                           ô
                                                                                                                                                                                                                                                                                                                                                       37
Claim 6; Fig 3; 71pp; English
                                                                                                                                                                                              100.0%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 51.8%;
100.0%;
                                                                                                                                                                                                                                                                                                                                                       Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        97us-0053035
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              98WO-US14547
                                                                                                                                                                                                                                                                                                                                                                                                              (first entry)
                                                                                                                                                                                                             Similarity 100.
37; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                           sednence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            1999-131752/11
                                                                                                                                                                                                                                                                                                                                                      AAY03207 standard;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             subtilis.
                                                                                                                                                                    56 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      37 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (UYMA-) UNIV
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      WO9903352-A1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            17-JUL-1998;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          18-JUL-1997;
                                                                                                                                                                                                                                                                                                                                                                                                              03-AUG-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  28-JAN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                          acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hansen JN;
                                                                                                                                                                                                                                                                                                                                                                                  AAY03207;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches' 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Bacillus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                      Sequence
                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     bacteria
                                                                                                                                        neutral
                                                                                                                                                                                                               Local
                                                                                                                                                                                                                                                                                                                                                                                                                                        Amino
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    nseq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            WPI;
                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Nev
                                                                                                                                                                                                                                                                                                                                         AAY03207
                                                                                                                                                                                                               Best
                                                                                                                                                                                                                                                                                                                         RESULT
```

XXXX

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        This is the amino acid sequence of sublancin 168 used in the method of the invention. The peptide designated sublancin 168, is an antimicrobial useful for treating infections and preserving food against spoilage bacteria, particularly Gram-positive bacteria. Pro-sublancin 168 and pre-sublancin 168, are the precursors of sublancin 168. Sublancin 168 is very stable at low pH and can be autoclaved without damage. It does not decompose after 2 years in aqueous solution of about neutral pH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                 New antimicrobial peptide, sublancin 168, from Bacillus subtilis used for, e.g. treatment of infections caused by Gram negative bacteria and as food preservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 37;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                               (Dha)"
                                                                                                                   (qua)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       .;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 96; DB 20
Pred. No. 1e-05;
); Mismatches
                                                                                 "dehydrogenated Ser
                                                                                                                     "dehydrogenated Thr
                                                                                                                                                      "thioether bridge"
Location/Qualifiers
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAG16276 standard; Protein; 408 AA.
                                                                                                                                                                                                                                                                                                                                          (UYMA-) UNIV MARYLAND BALTIMORE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ;
0
                                                                                                                                                                                                                                                                                                       97US-0053035.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   th 49.7%;
Similarity 94.4%;
17; Conservative
                                                                                                                                                                                                                                                                 98WO-US14547
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 GLGKAQCAALWLQCAXGG 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              20 GLGKAQCAALWLQCASGG
                   7..36
14..29
16
                                                                                                                 /note=
19..22
/note=
                                                                                 /note=
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Claim 13; Fig 7; 71pp;
                                                                                                                                                                                                                                                                                                                                                                                                                1999-131752/11.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   37 AA;
                          Disulfide-bond
Disulfide-bond
Modified-site
                                                                                                     Modified-site
                                                                                                                                                                                            W09903352-A1.
                                                                                                                                                                                                                               28-JAN-1999.
                                                                                                                                                                                                                                                                  17-JUL-1998;
                                                                                                                                                                                                                                                                                                       18-JUL-1997;
                                                                                                                                         Cross-links
                                                                                                                                                                                                                                                                                                                                                                              Hansen JN;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Seguence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAG16276
ID AAG1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
```

ö

ő

Gaps

ô

Indels

ed. No. 3. Mismatches

Pred

ő

Conservative

Similarity 18; Conser

.rag

```
990S-0139462.
990S-0139463.
990S-0139750.
990S-0139763.
990S-0139817.
990S-0139817.
990S-0140353.
990S-0140823.
990S-0144884.
990S-01448825.
990S-01448825.
990S-0144883.
990S-0144883.
990S-0144883.
990S-014488.
990S-0147303.
 pathway;
promoter;
                                          metabolic
n control;
                               16857
                                          n pathway; mexpression
                               ..
0N
                               11
                                          ; signal transduction enetic mapping; gene
                               SEQ
                               fragment
                              protein
                                                                                                               99US-0123180.
99US-0123180.
99US-0123548.
99US-0125788.
99US-0125788.
99US-0125785.
99US-0125785.
99US-0127462.
99US-0128234.
99US-0128234.
99US-0132486.
99US-0134219.
99US-0134218.
99US-013484.
99US-013484.
99US-013486.
99US-013486.
99US-013486.
99US-013484.
99US-0137502.
99US-0137502.
99US-0137502.
99US-0138847.
99US-01388847.
99US-01388847.
99US-0139455.
99US-0139455.
99US-0139455.
                                                                                                     -030143
                   entry)
                                         Protein identification hybridisation assay; gtermination sequence.
                              thaliana
                                                                  thaliana
                  (first
                                                                                                     2000EP
                                                                                                                  25-FEB-1999;
23-MAR-1999;
23-MAR-1999;
24-MAR-1999;
25-MAR-1999;
26-APR-1999;
27-APR-1999;
28-APR-1999;
29-MAY-1999;
21-APR-1999;
23-APR-1999;
24-MAY-1999;
25-MAY-1999;
26-MAY-1999;
27-MAY-1999;
27-MAY-1999;
28-APR-1999;
27-MAY-1999;
28-MAY-1999;
27-MAY-1999;
28-MAY-1999;
28-MAY-1999;
30-APR-1999;
30-AP
                                                                             EP1033405-A2
                                                                                                     -2000;
                  17-OCT-2000
                              Arabidopsis
                                                                  Arabidopsis
                                                                                         06-SEP-2000
      AAG16276;
                                                                                                     5-FEB
```

```
pathway;
promoter;
  metabolic control;
 n pathway; me
expression
  ; signal transduction enetic mapping; gene (
                                                                           990S-0121825.
990S-0123180.
990S-0123180.
990S-0123548.
990S-0125788.
990S-0125785.
990S-0126264.
990S-0126264.
990S-0128714.
990S-0130849.
990S-0134487.
990S-0134455.
990S-0134455.
990S-0139455.
990S-0139455.
990S-0139455.
990S-0139456.
990S-0139456.
990S-0139456.
990S-0139457.
                                                               2000EP-03014
 Protein identification hybridisation assay; gutermination sequence.
                          thaliana
                                                                5-FEB-2000;
                                       EP1033405-A2
                                                    06-SEP-2000.
                                                                          25-FEB-1999;
23-MAR-19999;
23-MAR-19999;
24-MAR-19999;
25-MAR-19999;
26-APR-19999;
27-APR-19999;
28-APR-19999;
29-MAY-19999;
21-APR-19999;
21-MAY-19999;
22-MAY-19999;
23-APR-19999;
24-MAY-19999;
25-MAY-19999;
26-MAY-19999;
27-MAY-19999;
28-MAY-19999;
28-MAY-19999;
29-MAY-19999;
30-APR-19999;
30-APR-19999;
31-APR-19999;
31-APR-19999;
31-APR-19999;
31-APR-19999;
32-APR-19999;
33-APR-19999;
33-APR-19999;
34-MAY-19999;
35-MAY-19999;
36-MAY-19999;
37-MAY-19999;
38-MAY-19999;
38-APR-19999;
39-MAY-19999;
39-MAY-19999;
31-AUN-19999;
31-AUN-19999;
31-AUN-19999;
31-AUN-19999;
31-AUN-19999;
31-AUN-19999;
33-AUN-19999;
33-AUN-19999;
33-AUN-19999;
                          Arabidopsis
  \alpha
                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                             9
                                                                                                                                                                                                                                                                                                                                                Length 408;
                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                          16856
                                                                                                                                                                                                                                                                                                                                                                                                                                                           ID NO:
                                                                                                                                                                                                                                                                                                                                                             10;
                                                                                                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                      268
                                                                                                                                                                                                                                                                                                                                                Score 58; DB Pred. No. 15; 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                          36
                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQ
                                                                                                                                                                                                                                                                                                                                                                                  ---WL-CLRG
                                                                                                                                                                                                                                                                                                                                                                         EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                          thaliana protein fragment
                                                                                                                                                                                                                                                                                                                                                                                                                     464
                                                                                                                                                                                                                                                                                                                                                            \ddot{5}
                                                                                                                                                                                                                                                                                                                                                                               ||| ||| ||| |:||||| EKLAKEVYF--LQNSEGSGLAPFDC
9905-0150566.
9905-0150884.
9905-0151065.
9905-0151080.
9905-0151080.
9905-0151438.
9905-0151438.
9905-0151438.
9905-0151438.
9905-0151438.
9905-0151438.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0160981.
9905-0160981.
9905-0160981.
9905-0160981.
                                                                                                                                                                                                                                                                                                                                                 .. ..
                                                                                                                                                                                                                                                                                                                                                                                                                     Protein;
                                                                                                                                                                                                                                                                                                                                               sh 30.18
Similarity 48.68
17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                               entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                             (first
                                                                                                                                                                                                                                                                                                                                                                                                                     standard;
25-AUG-1999;

27-AUG-1999;

27-AUG-1999;

27-AUG-1999;

31-AUG-1999;

31-AUG-1999;

31-AUG-1999;

31-AUG-1999;

31-AUG-1999;

31-AUG-1999;

32-SEP-1999;

33-OCT-1999;

34-OCT-1999;

34-OCT-1999;

34-OCT-1999;

35-OCT-1999;

37-OCT-1999;

38-OCT-1999;

38-OCT-1999;

38-OCT-1999;

38-OCT-1999;

38-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Arab!dopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                              -OCT-2000
                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG16275;
                                                                                                                                                                                                                                                                                                                                                                                                                    AAG16275
                                                                                                                                                                                                                                                                                                                                                                                     240
                                                                                                                                                                                                                                                                                                                                                                         2
                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 5
 a
                                                                                                                                                                                                                                                                                                                                                                                                                          XEXEX
                                                                                                                                                                                                                                                                                                                                                                                                                     a
                                                                                                                                                                                                                                                                                                                                                                         δ
```

.rag

```
99US-0140354.
99US-0140695.
99US-0140695.
99US-0141081.
99US-0141287.
99US-0141287.
99US-0141287.
99US-0142128.
99US-0142128.
99US-0142320.
99US-014232.
99US-014232.
99US-014433.
99US-014730.
99US-014433.
99US-014433.
99US-014433.
99US-014433.
99US-014730.
99US-014770.
99US-014933.
99US-014972.
99US-014972.
99US-014972.
99US-014972.
23 - JUN - 1999;

28 - JUN - 1999;

29 - JUN - 1999;

30 - JUN - 1999;

30 - JUN - 1999;

31 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

35 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

38 - JUL - 1999;

39 - JUL - 1999;

39 - JUL - 1999;

30 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

35 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

37 - JUL - 1999;

37 - JUL - 1999;

38 - JUL - 1999;

39 - JUL - 1999;

30 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

35 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

38 - JUL - 1999;

39 - JUL - 1999;

31 - JUL - 1999;

31 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

35 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

37 - JUL - 1999;

38 - JUL - 1999;

39 - JUL - 1999;

30 - JUL - 1999;

30 - JUL - 1999;

31 - JUL - 1999;

31 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

35 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

37 - JUL - 1999;

37 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1999;

33 - JUL - 1999;

34 - JUL - 1999;

36 - JUL - 1999;

37 - JUL - 1999;

38 - JUL - 1999;

39 - JUL - 1999;

30 - JUL - 1999;

31 - JUL - 1999;

32 - JUL - 1
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   crop; CBL; ENR-A; UROD;
cystathionine beta lyase;
porphobilinogen deaminase;
                                                                                                                                                                                                                                                                                                                                                                                               ø
                                                                                                                                                                                                                                                                                                                                                                                Length 464
                                                                                                                                                                                                                                                                                                                                                                                               10;
                                                                                                                                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                DB
18;
                                                                                                                                                                                                                                                                                                                                                                                         d. No. 18;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                        EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                Score 58;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Herbicide; porphobilinogen deaminase;
enoyl-acyl carrier protein reductase;
uroporphyrinogen decarboxylase; PBGD;
CPPO; coproporphyrinogen oxidase.
                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT 6
AAB47604
ID AAB47604 standard; Protein; 464
                                                                                                                                                                                                                                                                                                                                                                                                ;
;
                                                                                                                                                                                                                                                                                                                                                                                                                         CBL
99US-0152363.
99US-0152363.
99US-0153070.
99US-0153758.
99US-0154018.
99US-0154018.
99US-0154779.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0155486.
99US-0159231.
99US-0159231.
99US-0159293.
99US-0159294.
99US-0159331.
99US-0159637.
99US-0159638.
99US-0160767.
99US-0160981.
99US-0160981.
99US-0160981.
99US-0160989.
99US-0160989.
99US-0161406.
99US-0161360.
99US-0161360.
                                                                                                                                                                                                                                                                                                                                                                                th 30.1%;
Similarity 48.6%;
17; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       (first
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       beta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Cystathionine
 01-SEP-1999;

10-SEP-1999;

13-SEP-1999;

13-SEP-1999;

16-SEP-1999;

20-SEP-1999;

21-SEP-1999;

22-SEP-1999;

23-SEP-1999;

24-SEP-1999;

24-OCT-1999;

24-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

22-OCT-1999;

25-OCT-1999;

26-OCT-1999;

26-OCT-1999;

26-OCT-1999;

28-OCT-1999;

28-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -JAN-2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB47604;
                                                                                                                                                                                                                                                                                                                                                                                  y Match
Local S
                                                                                                                                                                                                                                                                                                                                                                                  Query Ma
Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                  7
                                                                                                                                                                                                                                                                                                                                                                                                                                 296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        07
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                   δ
```

thaliana

Arabidopsis

```
EP1033405-A2.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      25-FEB-1999;
05-MAR-1999;
09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
01-APR-1999;
06-APR-1999;
16-APR-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            25-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  06-SEP-2000.
                                                                                                                                                                        Disclosure;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          17-0CT-2000
                                           Levin JZ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AAG16274;
                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                  Sequence
                                                                      WPI; 200
N-PSDB;
 SAON)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            3;
                                                                                                                                                                                                                                                                                                                                         The sequences given in AAB47603-07 are proteins which may be used to identify herbicidal compounds, by combining one of these plant polypeptides, which have porphobilinogen deaminase activity, with a test compound under conditions allowing binding or inhibition. The method may be used for identifying herbicidal compounds, which can be used to suppress the growth of undesirable vegetation in fields where important crops are grown. The cDNA sequences encoding these proteins important crops are grown. The convergence of the plants of the convergence of the conv
                                                                                                                                                                                                                                                                        ndesirable
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            9
                                                                                                                                                                                                                                                                          in fields where important crops are grown, compria plant PBGD polypeptide with a test compound -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 464;
                                                                                                                                                                                                                                                                         of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                       Identifying herbicidal compounds for suppressing growth
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         herbicide tolerance;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     324
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 58; DB Pred. No. 18; 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  EKLAKEVYF -- LONSEGSGLAPFDC --- WL-CLRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                English
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              lyase
                                                                                                                                                                        AG
                                                                                                                                                                                                                                                                                                                              47pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          464
                                                                                                                                                                        PARTICIPATIONS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             cystathionine beta
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         CBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5
                                                                                                                                                                                                    Ŀ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               99US-0333366.
99US-0361879.
99US-0378313.
99US-0444117.
                                                                                  99US-198218P.
99US-240929P.
99US-228810P.
2000US-287572P.
2000US-287571P.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 48.6%;
Matches 17; Conservative
                                                                                                                                                                                                    Zheng
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        standard; Protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2000WO-EP05432
                                                         2000US-0586719
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  entry)
                                                                                                                                                                                                                                                                                                                               Column 55-58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cystathionine beta lyase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   thaliana.
                                                                                                                                                                                                    Bauer MW,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first
                                                                                                                                                                                                                               2001-638012/73
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     464 AA;
                                                                                                                                                                        SYNGENTA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      transgenic plant
                                                                                                                                                                                                                                            N-PSDB; AAH43576
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              WO200077185-A2
                                                                                    27 - JUL - 1999;
20 - AUG - 1999;
22 - NOV - 1999;
27 - MAR - 2000;
22 - MAY - 2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             15-JUN-1999;
27-JUL-1999;
20-AUG-1999;
22-NOV-1999;
01-DEC-1999;
                                                       05-JUN-2000;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13-JUN-2000;
US6294345-B1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Arabidopsis
                                                                                                                                                                                                                                                                                                                               Example 21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 28-MAR-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-DEC-2000
                            25-SEP-2001
                                                                                                                                                                                                                                                                                       vegetation
                                                                                                                                                                                                                                                                                                   combining
                                                                                                                                                                                                    Levin JZ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    AAB20017;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AAB20017
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           tissues
                                                                                                                                                                        (SYGN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AAB2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    du
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
```

```
The present sequence is that of Arabidopsis thaliana cystathionine beta lyase (CBL, EC-4.4.1.8), an enzyme that catalyses the conversion of cystathionine to homocysteine. The enzyme is essential for normal plant growth and development. The invention provides ENR-A, CBL, UROD, PBGD and CPPO genes (see AAA89291-95), methods for the recombinant production of the encoded enzymes (see AAB20016-20) in heterologous hosts, methods for enzymes centing themse herbicides to suppress growth of undesired vegetation. Plants, plant tissue, seeds and cells tolerant to herbicides that inhibit ENR-A, CBL, UROD, PBGD or CPPO are obtained by altering the corresponding enzyme activity either by increasing expression of wild-type herbicide-sensitive enzymes or by expressing modified herbicide tolerant enzymes. The herbicide
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.
                                                                                                                                                                                         New Arabidopsis thaliana DNA molecules for identifying compounds that are inhibitors of Arabidopsis enzymes encoded by the DNA, which can be used as herbicides -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 464;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Arabidopsis thaliana protein fragment SEQ ID NO: 16855.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     22;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 58; DB Pred. No. 18; 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      36
NOVARTIS AG.
NOVARTIS-ERFINDUNGEN VERW GES MBH.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                               English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 AAG16274 standard; Protein; 506 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  2;
                                                                                                                                                                                                                                                                                                93pp;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     99US-0121825.
99US-0123180.
99US-0123548.
99US-0125788.
99US-0126264.
99US-0127462.
99US-0128234.
99US-0128234.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  30.1%;
Similarity 48.6%;
17; Conservative
                                                                        Zheng
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000EP-0301439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (first entry)
                                                                                                                                                                                                                                                                                                  74-75;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Arabidopsis thaliana.
                                                                        Bauer MW,
                                                                                                                    2001-041382/05
B; AAA89292.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      464 AA;
                                                                                                                                                                                                                                                                                                     Page
```

3;

9US-014433 9US-0144435 9US-014488 9US-014481 9US-014508 9US-014508 9US-014508	9US-014514 9US-014521 9US-014522 9US-014527 9US-014591 9US-014591 9US-014591 9US-014591 9US-014591 9US-014595 9US-014730 9US-014730 9US-014730	9US-0147416 9US-0147493 9US-0147493 9US-0148171 9US-0148319 9US-0148341 9US-0148684 9US-0149426 9US-0149426 9US-0149426 9US-0149426 9US-0149929 9US-0149902 9US-0149902	990S-0151080. 990S-0151303. 990S-0151303. 990S-0151930. 990S-0153758. 990S-0154018. 990S-015486. 990S-0155486. 990S-0155659. 990S-0155659. 990S-0155659. 990S-015685. 990S-0157753. 990S-0157753. 990S-0157753.
9-JUL-1999 0-JUL-1999 0-JUL-1999 1-JUL-1999 1-JUL-1999 1-JUL-1999 2-JUL-1999 2-JUL-1999 2-JUL-1999	3-JUL-1999 3-JUL-1999 3-JUL-1999 6-JUL-1999 7-JUL-1999 7-JUL-1999 7-JUL-1999 2-AUG-1999 2-AUG-1999 4-AUG-1999 5-AUG-1999 6-AUG-1999	6-AUG-1999 9-AUG-1999 0-AUG-1999 0-AUG-1999 1-AUG-1999 3-AUG-1999 6-AUG-1999 6-AUG-1999 0-AUG-1999 0-AUG-1999 7-AUG-1999 6-AUG-1999 7-AUG-1999	27-AUG-1999; 30-AUG-1999; 31-AUG-1999; 01-SEP-1999; 10-SEP-1999; 13-SEP-1999; 15-SEP-1999; 20-SEP-1999; 22-SEP-1999; 24-SEP-1999; 24-SEP-1999; 24-SEP-1999; 04-OCT-1999; 05-OCT-1999; 06-OCT-1999; 13-OCT-1999; 13-OCT-1999; 13-OCT-1999; 13-OCT-1999;
PR PR PR PR PR	98 88 88 88 88 88 88 88 88 88 88 88 88 8	######################################	7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
9US-0130077 9US-0130449 9US-0130891 9US-0131449 9US-0132484 9US-0132484 9US-0132485 9US-0132485 9US-0132485	9US-0132863 9US-0134256 9US-0134218 9US-0134219 9US-0134370 9US-0134370 9US-0134368 9US-0135353 9US-0135629 9US-0135629 9US-013522 9US-0135353	9US-0137724 9US-0138094 9US-0138847 9US-0138847 9US-0139452 9US-0139453 9US-0139454 9US-0139455 9US-0139456 9US-0139460 9US-0139460 9US-0139460 9US-0139460 9US-0139460	990S-013981/. 990S-0139899. 990S-0140353. 990S-0140353. 990S-0140823. 990S-0141287. 990S-0142154. 990S-0142390. 990S-0142390. 990S-014232. 990S-0143542. 990S-0144331. 990S-0144331. 990S-0144331.
9-APR-1999 3-APR-1999 3-APR-1999 8-APR-1999 0-APR-1999 6-APR-1999 6-MAY-1999 6-MAY-1999	7-MAY-1999 1-MAY-1999 4-MAY-1999 4-MAY-1999 8-MAY-1999 9-MAY-1999 1-MAY-1999 1-MAY-1999 7-MAY-1999 3-JUN-1999	7-JUN-1999 8-JUN-1999 0-JUN-1999 6-JUN-1999 6-JUN-1999 8-JUN-1999 8-JUN-1999 8-JUN-1999 8-JUN-1999 8-JUN-1999 8-JUN-1999 8-JUN-1999	22-JUN-1999; 23-JUN-1999; 23-JUN-1999; 24-JUN-1999; 29-JUN-1999; 30-JUL-1999; 01-JUL-1999; 06-JUL-1999; 06-JUL-1999; 12-JUL-1999; 14-JUL-1999; 15-JUL-1999; 16-JUL-1999; 16-JUL-1999; 19-JUL-1999; 19-JUL-1999; 19-JUL-1999; 19-JUL-1999;
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~	**************************************	*

.rag

```
9905-0132484.
9905-0132486.
9905-0132486.
9905-0132486.
9905-0132487.
9905-0134218.
9905-0134219.
9905-0134211.
9905-0134211.
9905-0134211.
9905-0134211.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134221.
9905-0134231.
9905-0134231.
9905-0134231.
9905-0134232.
9905-0143622.
9905-0142390.
9905-0142390.
9905-0142390.
9905-0143431.
9905-0144333.
9905-0144333.
9905-0144333.
9905-0144333.
\mathbf{c}
                                                                                                                                                                             pathway;
promoter;
                                                                                                           Gaps
                                                                                                                                                                            Protein identification; signal transduction pathway; metabolic hybridisation assay; genetic mapping; gene expression control; termination sequence.
                                                                                                           9
                                                                                                   Length 506;
                                                                                                           Indels
                                                                                                                                                                     75196.
                                                                                                           10;
                                                                                                                                                                     ID NO:
                                                                                                   21;
                                                                                                                          366
                                                                                                                   36
                                                                                                  Score 58; DB
Pred. No. 19;
2; Mismatches
                                                                                                                                                                     SEQ
                                                                                                                  protein fragment
                                                                                                                                               ΑĄ
                                                                                                                                              200
                                                                                                           ;
;
                                                                                                  30.1%;
11 Similarity 48.6%;
17; Conservative
9905-0159330.
9905-0159331.
9905-0159638.
9905-0159584.
9905-0160741.
9905-0160770.
9905-0160814.
9905-0160815.
9905-0160980.
9905-0161404.
9905-0161406.
9905-0161406.
9905-0161406.
9905-0161359.
9905-0161359.
                                                                                                                                                                                                                           990S-0123180.
990S-0123180.
990S-0123548.
990S-0125788.
990S-0126264.
990S-0128234.
990S-0128714.
990S-0128714.
990S-0130891.
990S-0130891.
990S-0130891.
990S-0130891.
                                                                                                                                              standard; Protein;
                                                                                                                                                                                                                    2000EP-0301439
                                                                                                                                                              entry)
                                                                                                                                                                     thaliana
                                                                                                                                                                                             thaliana
                                                                                                                                                             (first
                                                                                                                                                                                                     EP1033405-A2
                                                                                                                                                                                                                    25-FEB-2000;
1-0CT-1999;
1-0CT-1999;
1-0CT-1999;
1-0CT-1999;
1-0CT-1999;
1-0CT-1999;
1-0CT-1999;
2-0CT-1999;
2-0CT-1999;
3-0CT-1999;
5-0CT-1999;
5-0CT-1999;
6-0CT-1999;
8-0CT-1999;
9-0CT-1999;
                                                                                                                                                                                                                           5-FEB-1999;

9-MAR-1999;

8-MAR-1999;

5-MAR-1999;

9-MAR-1999;

6-APR-1999;

6-APR-1999;

8-APR-1999;

1-APR-1999;

3-APR-1999;

3-APR-1999;

3-APR-1999;

3-APR-1999;

3-APR-1999;
                                                                                                                                                                                                            06-SEP-2000
                                                                                                                                                                     Arabidopsis
                                                                                                                                                                                             Arabidopsis
                                                                                                                                                             -OCT-2000
                                                                                                                                                      AAG58269;
                                                                                                   uery Match
                                                                                                                                              AAG58269
                                                                                                       Local
                                                                                                       Best Loc
Matches
                                                                                                                          338
                                                                                                                  ^{\circ}
                                                                                                                                      RESULT 9
 18
 d
                                                                                                                   Q
```

rag

٠

```
PR 22-JUL-1999 9908-0145085.

PR 22-JUL-1999 9908-01440087.

PR 22-JUL-1999 9908-01440087.

PR 23-JUL-1999 9908-01440087.

PR 23-JUL-1999 9908-0145218.

PR 27-JUL-1999 9908-0145218.

PR 27-JUL-1999 9908-0145218.

PR 27-JUL-1999 9908-0145218.

PR 27-JUL-1999 9908-0145218.

PR 02-AUG-1999 9908-0145918.

PR 03-AUG-1999 9908-014291.

PR 03-AUG-1999 9908-014291.

PR 03-AUG-1999 9908-014291.

PR 03-AUG-1999 9908-014308.

PR 03-AUG-1999 9908-014308.

PR 13-AUG-1999 9908-014308.

PR 23-AUG-1999 9908-014308.

PR 23-AUG-1999 9908-015308.

PR 23-AUG-1999 9908-015308.
```

```
pathway;
promoter;
                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                                                   pathway; metabolic expression control;
                                                                                                                                                      <u>~</u>
                                                                                                                                       Length 200;
                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                      thaliana protein fragment SEQ ID NO: 78966
                                                                                                                                                      10;
                                                                                                                                       21;
                                                                                                                                                                                                                                                                                   Protein identification; signal transduction hybridisation assay; genetic mapping; gene termination sequence.
                                                                                                                                       Score 56; DB Pred. No. 13; 5; Mismatches
                                                                                                                                        DB
                                                                                                                                                                    200 AA
                                                                                                                                                     5;
99US-0160768.

99US-0160770.

99US-0160815.

99US-0160981.

99US-0160981.

99US-0160989.

99US-0161404.

99US-0161405.

99US-0161405.

99US-0161360.

99US-0161359.

99US-0161359.

99US-0161920.

99US-0161992.
                                                                                                                                                                                                                                                                                                                                                                           990S-0121825.
990S-0123180.
990S-0123548.
990S-0126264.
990S-0126785.
990S-0127462.
990S-0127462.
990S-0128714.
990S-0130449.
990S-0130891.
990S-0131449.
990S-0132486.
990S-0132486.
990S-0132486.
990S-0132486.
990S-0132486.
                                                                                                                                                                                                                          standard; Protein;
                                                                                                                                       Match 29.0%;
Local Similarity 43.3%;
les 13; Conservative
                                                                                                                                                                                                                                                                                                                                                               2000EP-0301439
                                                                                                                                                                                                                                                        entry)
                                                                                                                                                                                                                                                                                                                  thaliana
                                                                                                                                                                                                                                                      (first
                                                                                                                                                                                                                                                                                                                                                                            25-FEB-1999;
09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
25-MAR-1999;
01-APR-1999;
06-APR-1999;
19-APR-1999;
21-APR-1999;
23-APR-1999;
23-APR-1999;
30-APR-1999;
30-APR-1999;
06-MAY-1999;
06-MAY-1999;
06-MAY-1999;
11-MAY-1999;
21-OCT-1999;
21-OCT-1999;
21-OCT-1999;
22-OCT-1999;
22-OCT-1999;
25-OCT-1999;
25-OCT-1999;
25-OCT-1999;
26-OCT-1999;
26-OCT-1999;
26-OCT-1999;
26-OCT-1999;
28-OCT-1999;
28-OCT-1999;
28-OCT-1999;
                                                                                                                                                                                                                                                                                                                                  EP1033405-A2
                                                                                                                                                                                                                                                                                                                                                               25-FEB-2000;
                                                                                                                                                                                                                                                                                                                                                06-SEP-2000.
                                                                                                                                                                                                                                                                                                                   Arabidopsis
                                                                                                                                                                                                                                                                      Arabidopsis
                                                                                                                                                                                                                                                        18-OCT-2000
                                                                                                                                                                                                                                        AAG60923;
                                                                                                                                       Query Match
Best Local S
Matches 13
                                                                                                                                                                                                                         AAG60923
                                                                                                                                                                                    128
                                                                                                                                                                                                           RESULT 10
AAG60923
                                                                                                                                                                      4
                                                                                                                                                                                                                   AAG609
                                                                                                                                                                                                                                                                              g
                                                                                                                                                                       ŏ
```

ij,

90S-014527 90S-014591 90S-014591 90S-014591 90S-014591 90S-014638 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720 90S-014720	990S-0149929. 990S-0149929. 990S-0149929. 990S-0149930. 990S-0150884. 990S-0151066. 990S-0151066. 990S-0151080. 990S-0151438. 990S-0151438. 990S-0151438. 990S-0151439. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0155486. 990S-0159330. 990S-0159294. 990S-0159293. 990S-0159293. 990S-0159293.
6-JUL-1999 7-JUL-1999 7-JUL-1999 8-JUL-1999 8-JUL-1999 8-JUL-1999 2-AUG-1999 4-AUG-1999 6-AUG-1999 6-AUG-1999 6-AUG-1999 6-AUG-1999 7-AUG-1999 7-AUG-1999 8-AUG-1999 9-AUG-1999 9-AUG-1999 9-AUG-1999 9-AUG-1999 9-AUG-1999 9-AUG-1999	23 - AUG - 1999; 25 - AUG - 1999; 25 - AUG - 1999; 26 - AUG - 1999; 27 - AUG - 1999; 27 - AUG - 1999; 31 - AUG - 1999; 32 - SEP - 1999; 33 - AUG - 1999; 34 - AUG - 1999; 35 - SEP - 1999; 36 - AUG - 1999; 37 - AUG - 1999; 38 - AUG - 1999; 39 - AUG - 1999; 30 - AUG - 1999; 31 - AUG - 1999; 31 - AUG - 1999; 32 - OCT - 1999; 31 - OCT - 1999; 32 - OCT - 1999; 33 - AUG - 1999; 34 - AUG - 1999; 55 - OCT - 1999; 56 - AUG - 1999; 57 - OCT - 1999; 57 - OCT - 1999; 57 - OCT - 1999; 58 - AUG - 1999; 59 - AUG - 1999; 70 - A
7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	**************************************
905-013422 905-013422 905-013437 905-013437 905-013437 905-013552 905-013535 905-013522 905-013752 905-013752 905-013752 905-013752 905-013752 905-013752 905-013752 905-013945 905-013945 905-013945	9905-0139459 9905-0139461 9905-0139462 9905-0139462 9905-0139462 9905-0139462 9905-0139463 9905-0139463 9905-0139463 9905-0139463 9905-0139463 9905-0144085 9905-0144085 9905-0144333 9905-0144333 9905-0144333 9905-0144333
4 - MAY - 1999 4 - MAY - 1999 8 - MAY - 1999 9 - MAY - 1999 1 - MAY - 1999 1 - MAY - 1999 2 - MAY - 1999 4 - JUN - 1999 3 - JUN - 1999 6 - JUN - 1999 6 - JUN - 1999 6 - JUN - 1999 8 - JUN - 1999	18 - JUN - 1999; 18 - JUN - 1999; 18 - JUN - 1999; 18 - JUN - 1999; 18 - JUN - 1999; 23 - JUN - 1999; 24 - JUN - 1999; 25 - JUN - 1999; 26 - JUN - 1999; 27 - JUN - 1999; 28 - JUN - 1999; 29 - JUN - 1999; 20 - JUL - 1999; 20 - JUL - 1999; 20 - JUL - 1999; 21 - JUL - 1999; 22 - JUL - 1999; 23 - JUL - 1999; 24 - JUL - 1999; 25 - JUL - 1999; 27 - JUL - 1999; 27 - JUL - 1999; 28 - JUL - 1999; 29 - JUL - 1999; 20 - JUL - 1999; 21 - JUL - 1999; 22 - JUL - 1999; 23 - JUL - 1999;
4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4	:

rag

~ 3

copy

3 1

-030

4

9 8

1 9

0 ı (C) ı

Thu

q

ŏ

```
990S-0135629.
990S-0135629.
990S-0135392.
990S-01377222.
990S-01377222.
990S-01377222.
990S-0137724.
990S-0137724.
990S-0137724.
990S-0137724.
990S-0137724.
990S-0138540.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139452.
990S-0139453.
990S-0140823.
990S-0143624.
990S-0140823.
990S-0143624.
990S-0143624.
990S-0144332.
990S-0144332.
990S-0144333.
990S-014433.
990S-0144332.
990S-0144333.
990S-0144333.
24 - MAY - 1999; 25 - MAY - 1999; 28 - MAY - 1999; 29 - JUN - 1999; 20 - J
    ij
                                                                                                                                                                                                                                                                                                                                                      pathway;
promoter;
                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                          O ..
                                                                                                                                                                                                                                                                                                                                                        metabolic
on control;
                                                                                                                                 Length 200;
                                                                                                                                                                                                                                                                                                                                                      expression contr
                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                  NO: 75195
                                                                                                                                                       10;
                                                                                                                                 21;
                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                         transduction
                                                                                                                               Score 56; DB
Pred. No. 13;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                  gene
                                                                                                                                                                                                                                                                                                                                   SEO
                                                                                                                                                                                                     157
                                                                                                                                                                              31
                                                                                                                                                                                                                                                                                                                                   thaliana protein fragment
                                                                                                                                                                                                                                                                                                                                                      Protein identification; signal transdesty bybridisation assay; genetic mapping; termination sequence.
                                                                                                                                                                                         LFKEVKLEELENQKGSGLGKAQCAAL - WL
                                                                                                                                                                                                                                                               Æ
                                                                                                                                                                                                                                                             202
                                                                                                                                                      5;
99US-0161404.
99US-0161405.
99US-0161359.
99US-0161360.
99US-0161361.
99US-0161920.
99US-0161992.
99US-0161993.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            990S-0121825.

990S-0123180.

990S-0125788.

990S-0126264.

990S-0126785.

990S-0126785.

990S-0128714.

990S-0128714.

990S-0130891.

990S-0130891.

990S-0132486.

990S-0132487.

990S-0132487.

990S-0132487.

990S-0132487.

990S-0132487.

990S-0132487.

990S-0134218.

990S-0134218.

990S-0134218.

990S-0134218.

990S-0134218.

990S-0134219.

990S-0134219.
                                                                                                                               th 29.0%;
Similarity 43.3%;
13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           2000EP-0301439
                                                                                                                                                                                                                                                             standard; Protein;
                                                                                                                                                                                                                                                                                                            entry)
                                                                                                                                                                                                                                                                                                                                                                                                       thaliana
                                                                                                                                                                                                                                                                                                           (first
25-OCT-1999;
25-OCT-1999;
26-OCT-1999;
26-OCT-1999;
26-OCT-1999;
28-OCT-1999;
28-OCT-1999;
28-OCT-1999;
28-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25-FEB-1999;
05-MAR-1999;
23-MAR-1999;
25-MAR-1999;
06-APR-1999;
06-APR-1999;
16-APR-1999;
16-APR-1999;
23-APR-1999;
23-APR-1999;
30-APR-1999;
30-APR-1999;
30-APR-1999;
06-MAY-1999;
06-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
19-MAY-1999;
19-MAY-1999;
19-MAY-1999;
19-MAY-1999;
19-MAY-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                            EP1033405-A2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          25-FEB-2000;
                                                                                                                                                                                                                                                                                                          18-OCT-2000
                                                                                                                                                                                                                                                                                                                                                                                                       Arabidopsis
                                                                                                                                                                                                                                                                                                                                   Arabidopsis
                                                                                                                                                                                                                                                                                                                                                                                                                                                   -SEP-2000
                                                                                                                               Query Match
Best Local S
Matches 13
                                                                                                                                                                                                                                                            AAG58268
                                                                                                                                                                                                                                                                                   AAG58268
                                                                                                                                                                                                                                      RESULT 11
AAG58268
                                                                                                                                                                              4
                                                                                                                                                                                                     128
                                                                                                                                                                                                                                                                                                                                                                                                                                                   90
```

ત્ને

.rag

```
PR 02-NuG-1999; 99us-0146389.
PR 03-AuG-1999; 99us-0147304
PR 04-AuG-1999; 99us-0147304
PR 06-AuG-1999; 99us-0147304
PR 06-AuG-1999; 99us-0147302
PR 06-AuG-1999; 99us-0147303.
PR 10-AuG-1999; 99us-0147416.
PR 11-AuG-1999; 99us-0144146.
PR 11-AuG-1999; 99us-0144173.
PR 11-AuG-1999; 99us-0144173.
PR 12-AuG-1999; 99us-0144173.
PR 12-AuG-1999; 99us-0144173.
PR 12-AuG-1999; 99us-0144173.
PR 22-AuG-1999; 99us-015468.
PR 22-AuG-1999; 99us-015468.
PR 22-AuG-1999; 99us-015468.
PR 22-SEP-1999; 99us-015518.
PR 22-SEP-1999; 99us-015518.
PR 23-AuG-1999; 99us-015518.
PR 24-SEP-1999; 99us-015518.
PR 25-AuG-1999; 99us-015518.
PR 21-CCT-1999; 99us-015518.
PR 11-CCT-1999; 99us-015529.
PR 12-CCT-1999; 99us-015529.
PR 13-CCT-1999; 99us-015529
```

```
il transduction pathway; metabolic pathway; mapping; gene expression control; promoter;
                                             Gaps
                                           2;
                             Length 202;
                                             Indels
                                                                                                                                                            thaliana protein fragment SEQ ID NO: 78965
                                            10;
                              21;
                             Score 56; DB Pred. No. 13; Mismatches
                                                                          15
                                                           31
                                                                   LFKEVKLEELENQKGSGLGKAQCAAL--WL
                                                                                                               202
                                           5.
                                                                                                                                                                         Protein identification; signal hybridisation assay; genetic matermination sequence.
                                                                                                                                                                                                                                                                 990s-0123180.
990s-0123180.
990s-0123548.
990s-0126264.
990s-0126785.
990s-0126785.
990s-0128714.
990s-0130449.
990s-0130449.
990s-0132486.
990s-0132487.
990s-0132487.
990s-0132487.
990s-0132487.
990s-0132487.
990s-0134219.
99US-0161993.
99US-0161993.
99US-0162142.
                            Query Match
Best Local Similarity 43.3%;
Matches 13; Conservative
                                                                                                               standard; Protein;
                                                                                                                                                                                                                                                    2000EP-0301439
                                                                                                                                             entry)
                                                                                                                                                                                                        Arabidopsis thaliana
                                                                                                                                            (first
 \sigma
                                                                                                                                                                                                                                                                25-FEB-1999;
09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
01-APR-1999;
06-APR-1999;
16-APR-1999;
16-APR-1999;
23-APR-1999;
23-APR-1999;
30-APR-1999;
30-APR-1999;
30-APR-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
18-MAY-1999;
20-MAY-1999;
21-MAY-1999;
21-MAY-1999;
22-MAY-1999;
24-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
27-MAY-1999;
28-OCT-1999;
28-OCT-1999;
29-OCT-1999;
                                                                                                                                                                                                                      EP1033405-A2
                                                                                                                                                                                                                                                    25-FEB-2000;
                                                                                                                                                           Arabidopsis
                                                                                                                                                                                                                                     06-SEP-2000
                                                                                                                                            18-OCT-2000
                                                                                                              AAG60922
                                                                                                                              AAG60922
                                                                         130
                                                            4
                                                                                                 12
                                                                                                RESULT 12
AAG60922
                                                                                                                      PR
PR
PR
                                                                          g
                                                            δý
```

```
990S-0147416.
990S-0147493.
990S-0147493.
990S-0148171.
990S-0148319.
990S-0148341.
990S-0148341.
990S-0148341.
990S-0148565.
990S-0148565.
990S-0158029.
990S-0157753.
990S-0158029.
990S-0157753.
990S-0158029.
990S-015838.
990S-015838.
990S-015939.
990S-015939.
990S-015939.
990S-015939.
990S-015939.
990S-015939.
990S-015929.
990S-015939.
990S-0159884.
990S-015929.
990S-0159884.
990S-0159884.
990S-0159884.
990S-0159889.
990S-0159884.
990S-0159884.
990S-0159884.
990S-0159884.
990S-0159889.
990S-0159884.
990S-0159884.
990S-0159888.
06-AUG-1999;

09-AUG-1999;

10-AUG-1999;

11-AUG-1999;

12-AUG-1999;

13-AUG-1999;

13-AUG-1999;

13-AUG-1999;

14-AUG-1999;

20-AUG-1999;

21-AUG-1999;

20-AUG-1999;

21-AUG-1999;

22-AUG-1999;

22-AUG-1999;

23-AUG-1999;

24-SEP-1999;

25-AUG-1999;

27-AUG-1999;

28-SEP-1999;

29-SEP-1999;

21-OCT-1999;

21-OCT-1999;

22-SEP-1999;

23-AUG-1999;

24-SEP-1999;

25-AUG-1999;

26-AUG-1999;

27-AUG-1999;

27-AUG-1999;

28-SEP-1999;

29-SEP-1999;

21-OCT-1999;

21-OCT-1999;

21-OCT-1999;

22-OCT-1999;

22-OCT-1999;

23-OCT-1999;

24-OCT-1999;

25-OCT-1999;

26-OCT-1999;

26-OCT-1999;

26-OCT-1999;

26-OCT-1999;

26-OCT-1999;

26-OCT-1999;

27-OCT-1999;

28-OCT-1999;

28-OCT-1999;

28-OCT-1999;

28-OCT-1999;

28-OCT-1999;

28-OCT-1999;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Ouery Match
Best Local S
Matches 13
   990S-0137724

990S-0138640.

990S-0138540.

990S-0139847.

990S-0139452.

990S-0139452.

990S-0139452.

990S-0139457.

990S-0139457.

990S-0139457.

990S-014085.

990S-0144085.

990S-0144085.

990S-0144333.

990S-0144335.

990S-0144332.

990S-0144333.

990S-0144333.

990S-0144332.

990S-0144332.

990S-0144332.

990S-0144333.

990S-0144333.

990S-0144333.

990S-0144333.

990S-0144333.

990S-0145089.

990S-0145089.

990S-0145089.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.

990S-0145318.
  02-70N-19
10-70N-19
116-70N-19
116-70N-19
118-70N-19
123-70N-19
133-70N-19
14-70N-19
15-70N-19
16-70N-19
17-70N-19
18-70N-19
18-70N-19
18-70N-19
18-70N-19
18-70N-19
19-70N-19
19
```

Query Match

29.0%; Score 56; DB 21; Length 202;

Best Local Similarity 43.3%; Pred. No. 13;

Matches 13; Conservative 5; Mismatches 10; Indels 2;

1;

Gaps

Ħ

```
990S-0139454.
990S-0139455.
990S-0139455.
990S-0139455.
990S-0139455.
990S-0139455.
990S-0139460.
990S-0139460.
990S-0139461.
990S-0139461.
990S-0139463.
990S-0144332.
990S-0144332.
990S-0144332.
990S-0144332.
990S-0144333.
pathway;
promoter;
                                                             U ..
                                                             ---
                                                         riviein identification; signal transduction pathway; metabol. hybridisation assay; genetic mapping; gene expression controtermination sequence.
                                                     78964
                                                     thaliana protein fragment SEQ ID NO:
            159
        LFKEVKLEELENQKGSGLGKAQCAAL - WL
                               AA
                              215
                                                                                                        9905-0121825.
9905-0123180.
9905-0123548.
9905-0125788.
9905-0126264.
9905-0126264.
9905-0126264.
9905-0130891.
9905-0132487.
9905-0132486.
9905-0132486.
9905-0132486.
9905-0132487.
9905-0132487.
9905-0132486.
9905-0132486.
9905-0132486.
9905-0132486.
9905-0132487.
9905-0132486.
9905-0134218.
9905-0134218.
9905-0135529.
9905-0135529.
9905-0135529.
9905-0135528.
9905-0135528.
9905-0137528.
9905-0137528.
                              AAG60921 standard; Protein;
                                                                                                  2000EP-0301439
                                              entry)
                                                                            Arabidopsis thaliana
                                             (first
                                                                                                         25-FEB-1999;
09-MAR-1999;
23-MAR-1999;
25-MAR-1999;
01-APR-1999;
06-APR-1999;
16-APR-1999;
16-APR-1999;
16-APR-1999;
30-APR-1999;
30-APR-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
15-MAY-1999;
16-JUN-1999;
01-JUN-1999;
01-JUN-1999;
01-JUN-1999;
01-JUN-1999;
01-JUN-1999;
16-JUN-1999;
16-JUN-1999;
16-JUN-1999;
16-JUN-1999;
                                                                                   EP1033405-A2
                                                                                                  25-FEB-2000;
                                                     Arabidopsis
                                             18-OCT-2000
                                                                                           -SEP-2000
                                      AAG60921
                      RESULT 13
AAG60921
           130
    4
                                                                                           90
                                  a
    ð
```

```
transduction pathway; metabolic pathway; ppping; gene expression control; promoter;
                                           ID NO: 75194
                                           thaliana protein fragment
                                                              mapping
        AA
       standard; Protein; 216
                                                      Protein identification; signal hybridisation assay; genetic metermination sequence.
                                                                                                                               99US-0121825.

99US-0123180.

99US-0123548.

99US-0126264.

99US-0126264.

99US-0126785.

99US-0126785.

99US-0128714.

99US-0132487.

99US-0132487.

99US-0132486.

99US-0132486.

99US-0132486.

99US-0132486.

99US-0132486.

99US-0132486.

99US-0132487.

99US-0132486.

99US-0132487.

99US-0134218.

99US-0134218.

99US-0137222.

99US-0137222.

99US-0137222.

99US-0137222.

99US-0137222.

99US-0137358.

99US-0137324.

99US-0137324.

99US-0137324.

99US-0137322.

99US-0137322.

99US-0139452.

99US-0139452.

99US-0139455.

99US-0139455.

99US-0139455.
                                                                                                                    2000EP-0301439
                                entry)
                                                                                thaliana
                               (first
                                                                                                                               25-FEB-1999;
09-MAR-1999;
23-MAR-1999;
23-MAR-1999;
25-MAR-1999;
01-APR-1999;
06-APR-1999;
16-APR-1999;
16-APR-1999;
23-APR-1999;
23-APR-1999;
24-MAY-1999;
14-MAY-1999;
06-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
14-MAY-1999;
15-MAY-1999;
16-MAY-1999;
16-MAY-1999;
16-MAY-1999;
16-MAY-1999;
16-JUN-1999;
16-JUN-1999;
16-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
18-JUN-1999;
                                                                                            EP1033405-A2
                                                                                                                    25-FEB-2000;
                                                                                Arabidopsis
                                                                                                         06-SEP-2000.
                                            Arabidopsis
                               18-OCT-2000
AAG58267
ID AAG58267
              ï
                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                            7
                                                                                                                                                                                                                                                                                                                                                                                                21
                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                            10;
                                                                                                                                                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                               Score 56; DB
Pred. No. 14;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                    17
                                                                                                                                                                                                                                                                                                                                                                                                                        31
                                                                                                                                                                                                                                                                                                                                                                                                                               -WL
                                                                                                                                                                                                                                                                                                                                                                                                                        LFKEVKLEELENOKGSGLGKAQCAAL
                                                                                                                                                                                                                                                                                                                                                                                                            S
99US-0148684.
99US-0149368.
99US-0149426.
99US-0149426.
99US-0149422.
99US-0149723.
99US-0149929.
99US-0149929.
99US-0150884.
99US-0150884.
99US-0151080.
99US-0151333.
99US-0151333.
99US-0151333.
99US-0151333.
99US-0151333.
99US-0151333.
99US-0159294.
99US-0159293.
99US-0160981.
99US-0160981.
                                                                                                                                                                                                                                                                                                                                                                                               Query Match
Best Local Similarity 43.3%;
Matches 13; Conservative
```

-OCT-19

200

4

Q

14

d

RESULT

13-Aug-19
16-Aug-19
17-Aug-19
18-Aug-19
20-Aug-19
20-Aug-19
20-Aug-19
23-Aug-19
23-Aug-19
24-SEP-19
25-Aug-19
26-Aug-19
27-Aug-19
27-Aug-19
28-SEP-19

.rag

```
RR 18-70N-1999, 9918-013946.
RR 18-70N-1999, 9918-013946.
RR 18-70N-1999, 9918-013946.
RR 22-70N-1999, 9918-013946.
RR 23-70N-1999, 9918-013946.
RR 23-70N-1999, 9918-014184.
RR 19-70L-1999, 9918-01418.
RR 22-70L-1999, 9918-01418.
RR 22-70L-1
```

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 216
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                21;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DB
14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 56; DB Pred. No. 14; 5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           70
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        5;
9905-0149902.
9905-0149930.
9905-0150566.
9905-0150884.
9905-0151086.
9905-0151086.
9905-0151080.
9905-0151080.
9905-0151080.
9905-0151080.
9905-0151080.
9905-0151080.
9905-0151333.
9905-0151333.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159293.
9905-0159638.
9905-0159638.
9905-015929.
9905-0159638.
9905-0159638.
9905-0159638.
9905-0159638.
9905-0160768.
9905-016098.
9905-016098.
9905-016098.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           standard; peptide;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 43.38, les 13; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 entry)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               (first
23-AUG-1999;
25-AUG-1999;
26-AUG-1999;
27-AUG-1999;
27-AUG-1999;
30-AUG-1999;
31-AUG-1999;
31-AU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              13-SEP-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ST 15
1125
AAY24125
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AAY24125;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best_Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             144
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  uery
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT AAY2417 ID A XX AC XX XX DT XX
    qq
```

1;

```
<del>;</del>
                                                                                                                                                                                                                                                                                                                                                                                                             The present invention describes gamma-carboxylated conopeptides derived from cone snail venom. The gamma-conopeptides and their propeptides are useful as agonists of neuronal pacemaker calcium channels. The conopeptides are naturally available in minute amounts in the venom of cone snails and their derivatives are synthetic. The peptides modulate slow inward cation channels in vertebrates involved in syndromes of clinical relevance, such as epileptic activity in hippocampus and pacemaker potentials in heart muscle. The present sequence represents a conopeptide propeptide.
                        Conopeptide; gamma-conopeptide; venom; cone snail; cation channel; epilepsy; pacemaker; heart muscle; neuronal pacemaker calcium channel.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                Gamma-carboxylated conopeptides used as, e.g. neuronal pacemaker
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 70;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                 Walker C;
                                                                                                                                                                                                                                                  Burlingame AL, Colledge C, Cruz LJ, Fainzilber M;
Imperial J, Kits KS, Olivera BM, Shetty R, Walke
Walkins M;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2 EKLFKEVKLEELENQKGSGLGKAQ-----CAALWLQCASG 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         57
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB 20;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               OALYQEKRKEMINFLSKGKINAERRNEGCKATWMSCSSG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 55.5; DB Pred. No. 5.1; 9; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                       Claim 37; Page 52; 61pp; English.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9;
Conopeptide propeptide Mr6.3
                                                                                                                                                                     97US-0069706.
                                                                                                                                           98WO-US26792.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28.8%;
30.0%;
                                                                                                                                                                                                CALIFORNIA.
UTAH RES FOUND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local Similarity
                                                                                                                                                                                                                                                                                                       WPI; 1999-418708/35.
N-PSDB; AAX88505.
                                                                                                                                                                                                                         (UYVR-) UNIV VRIJE.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 70 AA;
                                                                                                                                                                                                                                                                                                                                                            calcium channels
                                                                 marmoreus.
                                                                                        WO9930732-A1.
                                                                                                                                                                                                UNIA
                                                                                                                                            16-DEC-1998;
                                                                                                                                                                       16-DEC-1997;
                                                                                                                 24-JUN-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence
                                                                                                                                                                                                (REGC
(UTAH
                                                                Conus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         18
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         q
```

09:28:00

Search completed: October 24, 2002, Job time: 16.3139 secs

		42.	
			•
			•

Appli

4, Appli 21, Appl 21, Appl 56, Appl 56, Appl 59, Appl 57, Appl 57, Appl 57, Appl 6, Appli 6, Appli 6, Appli 6, Appli 6, Appli 6, Appli 3, Appli 3, Appli 3, Appli 3, Appli

Sequence Sequence

```
Length 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           and Methods
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           10;
PCT-US94-09752-4
US-08-145-995A-21
US-08-451-747-21
US-09-134-852-21
US-08-325-071-59
US-08-461-004A-56
US-08-461-004A-59
US-08-461-004A-59
US-08-461-004A-57
US-08-461-004A-57
US-08-461-004A-57
US-08-319-554A-6
US-08-319-554A-6
US-08-319-554A-6
US-08-315-6
US-08-319-554A-6
US-08-315-071-69
US-08-325-071-69
US-08-325-071-69
US-08-325-071-69
US-08-325-071-69
US-08-325-071-69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            4;
                                                                                                                                                                                                                                                                                                                                                                                                                  RESULT 1
US-09-586-719-4

Sequence 4, Application US/09586719

Patent No. 6294345

GENERAL INFORMATION:

APPLICANT: Zheng, Feng

APPLICANT: Levin, Joshua Z.

APPLICANT: Bauer, Michael W.

TITLE OF INVENTION: Herbicide Target Genes and FILE REFERENCE: PB/5-30909A

CURRENT APPLICATION NUMBER: US/09/586,719

CURRENT FILING DATE: 2000-06-05

NUMBER OF SEQ ID NOS: 34

SOFTWARE: PatentIn Ver. 2.1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 58; DB Pred. No. 4.4; 2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ••
                                                                                                                                                                                                                                                                                                                                                              ALIGNMENTS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NOVEL COMPOUNDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 2
US-09-197-636-2
US-09-197-636-2
; Sequence 2, Application US/09197636
; Patent No. 6239267
; GENERAL INFORMATION:
    APPLICANT: DUCKWORTH, DAVID
    APPLICANT: HAYES, PHILIP
    APPLICANT: MEADOWS, HELEN
    APPLICANT: MEADOWS, HELEN

                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COUNTRY: 0.5
ZIP: 19482-0980
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ORGANISM: Arabidopsis thaliana
S-09-586-719-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            30.1%;
illarity 48.6%;
Conservative
         284
591
5591
650
650
650
688
688
828
82
82
147
147
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              1 Similarity
17; Conserv
       ID NO 4
FNGTH: 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Query Match
Best Local S
Matches 17
          PRT
                                                                                                                                                                                                                                                                                            44
                                                                                                                                                                                    45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ~
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TYPE:
          g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            34, Appl
61, Appl
65, Appl
65, Appl
63, Appl
67, Appl
67, Appl
8, Appli
3, Appli
3, Appli
3, Appli
2, Appli
3, Appli
2, Appli
6, Appli
7, Appli
8, Appli
8, Appli
8, Appli
7, Appli
8, Appli
8, Appli
8, Appli
8, Appli
9, Appli
10, Appli
10, Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
Appli
Appli
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Appli
                                                                                                                                                                       updates/sec
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ing printed
                                                                                                                                        Seconds
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ription
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Sequence
Seq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 have
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Desci
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               predicted by chance to
score of the result be
total score distribution
                                                                                                                                    ; Search time 5.6715.
(without alignments)
159.348 Million cell
                                                                                                                                                                                                                                                                                                                                                                        231628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     'ued_Patents_AA:*
/cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
/cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
/cgn2_6/ptodata/1/iaa/bacyCOMB.pep:*
/cgn2_6/ptodata/1/iaa/backfiles1.pep:*
                                                    Ltd
                                                                                                                                                                                                                                                 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                5.1.3
Compugen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-09-586-719-4
US-09-197-636-2
US-09-197-636-4
US-09-197-636-8
US-09-197-636-8
US-09-197-636-8
US-09-235-451-34
US-08-325-071-61
US-08-325-071-65
US-08-325-071-65
US-08-461-004A-65
US-08-325-071-67
US-08-461-004A-67
US-08-461-004A-67
US-08-461-004A-67
US-08-461-004A-8
US-08-471-491-3
US-08-118-319-5
US-09-118-319-5
US-08-254-989-2
US-08-981-392-2
US-08-981-392-2
US-08-872-855-8
US-09-220-528-104
                                                                                                                                                                                                                                                                                                                                                                           chosen parameters
                                                                                                                                                                                                                                                                                                                                        residues
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SUMMARIES
                                version - 2002
                                                                                                                                                                                                                                                                                                                                                                                                                                                              : Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
                                                                                                       model
                                                                                                                                        :01
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                results
1 to the
of the t
                                                                                                                                         09:27
                                                                                                                                                                                                              _COPY_1
                                                                                                                                                                                                                                                                                                      0
                                                                                                                                                                                                                                                                                                                                     seds, 24425594
                                                                                                         ΝS
                                 GenCore
(c) 1993
                                                                                                                                                                                                                                                                                                       Gapext
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    n or equal analysis or
                                                                                                         using
                                                                                                                                                                                                                                                                                                                                                                                                             2000000000
                                                                                                                                           2002,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               よみよなこなこなましななころのなけるないのろろろろすろみ
                                                                                                                                                                                                        US-09-894-030-3_0
193
1 MEKLFKEVKLEET.F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            DB
                                                                                                                                                                                                                                                                                                                                                                             satisfying
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  the number
than or eq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            464
839
839
839
839
839
839
620
650
650
650
1296
1296
1287
1287
1287
1287
1287
1287
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length
                                                                                                                                                                                                                                                                                                            •
                                                                                                         search,
                                                                                                                                           24,
                                                      Copyright
                                                                                                                                                                                                                                                                                      BLOSUM62
Gapop 10.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    No. is the nigreater than s derived by
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Issued_
                                                                                                                                                                                                                                                                                                                                                                                                               length:
length:
                                                                                                                                             October
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           231628
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          %
Query
Match
                                                                                                                                                                                                                                                                                                                                                                             hits
                                                                                                         protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              .. .. .. .. .. ..
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            H 2 M 4 5 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            5.54

5.44

5.45

5.45

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

5.51

                                                                                                                                                                                                                                                                                                                                                                             of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Post-processing
                                                                                                                                                                                                                                                                                                                                                                                                                 sed
                                                                                                                                                                                                                                                                                       table:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Score
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        score gand is
                                                                                                                                                                                                                                     score
                                                                                                                                                                                                                                                                                                                                                                             number
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Pred
                                                                                                                                                                                                                                                                                                                                                                                                                DB
DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ••
                                                                                                          protein
                                                                                                                                                                                                                                                      Sequence:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Database
                                                                                                                                                                                                                                                                                                                                           Searched
                                                                                                                                                                                                                   Title:
Perfect
                                                                                                                                                                                                                                                                                                                                                                                                                  Minimum
                                                                                                                                                                                                                                                                                                                                                                                                                                     Maximum
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 10m450/86010m450/20010m450/
                                                                                                                                                                                                                                                                                          Scoring
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAAAAAAAAAAAAAAAAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Result
No.
                                                                                                                                              Run on
                                                                                                                                                                                                                                                                                                                                                                              Total
```

3

Gaps

9

Indels

```
28.2%;
larity 43.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      IBM Compatible
                                                                                                                                          .. 639 amino acids
amino acid
       INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        839 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TELEX: 846169
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZIP: 19482-0980
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
                                                                                                                                                                                      single
                                                                                                                                                                                                                   protein
                                                                                                                                                                                                                                                                                                                           KLFKEVKLEELENOKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       amino acid
                                                                                                                                                                                                     linear
                                                                                                                                                                                                                                                                            Best Local Similarity
Matches 16; Conser
                                                                                                                                                                       TYPE: amino a STRANDEDNESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TYPE: amino (STRANDEDNESS:
                                                                                                                                                                                                   ; TOPOLOGY; II; MOLECULE TYPE:
US-09-197-636-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             \OmegaS
      ATTORNEY/AGENT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       COMPUTER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                           m
                                                                                                                                                                                                                                                                                                                                                       QQ
                                                                                                                                                                                                                                                                                                                             ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,636
FILING DATE: 23-NOV-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 54.5; D Pred. No. 24; 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       -SGLGKAQCAALWLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       for Windows Version
                                                                                                                    FILING DATE: 12-MAR-1998
APPLICATION NUMBER: UK 9815791.0
FILING DATE: 21-JUL-1998
APPLICATION NUMBER: UK 9819278.4
FILING DATE: 03-SEP-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30075
TELECOMMUNICATION INFORMATION:
TELEPHONE: 601-407-0701
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Ver:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,636
FILING DATE: 23-NOV-1998
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   COMPOUNDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 12-MAR-1998
APPLICATION NUMBER: UK 9815791.0
FILING DATE: 21-JUL-1998
APPLICATION NUMBER: UK 9819278.4
FILING DATE: 03-SEP-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            UK 9805137.8
                                                                                      PRIOR APPLICATION DATA:
APPLICATION NUMBER: UK 9805137
FILING DATE: 12-MAR-1998
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Sequence 4, Application US/09197636 Patent No. 6239267 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         3,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Prestia
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STATE: PA
COUNTRY: US
ZIP: 19482-0980
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DUCKWORTH, DAVID HAYES, PHILIP MEADOWS, HELEN DAVIS, JOHN
                                                                                                                                                                                                                                                                                                                                                                                                                                                            28.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NOVEL
                                                                                                                                                                                                                                                                                                                                                       amino acids
                                                                                                                                                                                                                                                                                                                   SEQUENCE CHARACTERISTICS
LENGTH: 839 amin
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & F
STREET: P.O. Box 980
CITY: Valley Forge
                                                                                                                                                                                                                                                                                                                                                                                    single
                                                                                                                                                                                                                                                                                                                                                                                                                 protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICANT: DUCKWORTH,
APPLICANT: HAYES, PHII
APPLICANT: MEADOWS, HE
APPLICANT: DAVIS, JOHN
TITLE OF INVENTION: NC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3 KLFKEVKLEELENOKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                  amino acid
                                                                                                                                                                                                                                                                                                                                                                                                   linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Similarity
16; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               FILING DATE: 2 CLASSIFICATION:
                                                                                                                                                                                                                                                                                                          TELEX: 846169 INFORMATION FOR SE
                                                                                                                                                                                                                                                                                                                                                                               STRANDEDNESS:
; TOPOLOGY: line
; MOLECULE TYPE: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT 3
US-09-197-636-4
                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       à
```

```
Gaps
                                                                                                                                                                                                                                                                                                                    7;
                                                                                                                                                                                                                                                                                    Length 839;
                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                      11;
                                                                                                                                                                                                                                                                                      DB 4;
                                                                                                                                                                                                                                                                                                                                                                                            355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                    Score 54.5; Di
Pred. No. 24;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                             319 KLHPTLKLEELTNKKGMTPLALAAGTGKIGVLAYILO
                                                                                                                                                                                                                                                                                                                                                           -SGLGKAQCAALWLQ
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30075
TELECOMMUNICATION INFORMATION:
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 839 amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:

APPLICATION NUMBER: UK 9805137.8
FILING DATE: 12-MAR-1998
APPLICATION NUMBER: UK 9815791.0
FILING DATE: 21-JUL-1998
APPLICATION NUMBER: UK 9819278.4
FILING DATE: 03-SEP-1998
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GP-30075
TELEPHONE: 601-407-0701
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICATION DATA:
FILING DATE: 23-NOV-1998
CLASSIFICATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      NOVEL COMPOUNDS
                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT 4
US-09-197-636-8
; Sequence 8, Application US/09197636
; Patent No. 6239267
; GENERAL INFORMATION:
; APPLICANT: DUCKWORTH, DAVID
                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    STREET: P.O. BOX 980
CITY: Valley Forge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: HAYES, PHILIP
APPLICANT: MEADOWS, HELEN
APPLICANT: DAVIS, JOHN
TITLE OF INVENTION: NOVEL C(
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rather & Drocket
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ..
œ
```

ί

#

Version

```
ZIP: 20007-5109
COMPUTER READABLE FORM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
hes 10; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ; TYPE: amino a
; TOPOLOGY: lis
; MOLECULE TYPE:
US-08-325-071-61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             US-08-461-004A-61; Sequence 61, Ap
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ij
                                                                                        ...
....
                                                                                                                                                                                                                                               APPLICANT: Julius, David J.
APPLICANT: Caterina, Michael J.
APPLICANT: Caterina, Michael J.
APPLICANT: Brake, Anthony J.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES ENCODING
TITLE OF INVENTION: CAPSAICIN RECEPTOR AND CAPSAICIN RECEPTOR-RELATED
TITLE OF INVENTION: POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: 9076/084CIP
CURRENT APPLICATION NUMBER: US/09/235,451
CURRENT FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/072,151
PRIOR FILING DATE: 1998-01-22
PRIOR FILING DATE: 1997-08-20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             6
                                                          Length 839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Length 83
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Membrane
                                                                                     11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB 4;
                                                          DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                355
                                                                                                                                               355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32
                                                                                                                    32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Men
TITLE OF INVENTION: Glycoprotein Of A Tick
NUMBER OF SEQUENCES: 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 54.5; D
pred. No. 24;
3; Mismatches
                                                         Score 54.5; D
Pred. No. 24;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -SGLGKAQCAALWLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -SGLGKAQCAALWLQ
                                                                                                                                  PRIOR FILING DATE: 1997-08-20
NUMBER OF SEQ ID NOS: 48
SOFTWARE: FastSEQ for Windows Version 3.0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Stewart Gary
Joanna Terry
V, Law Anthony York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Sequence 61, Application US/08325071 Patent No. 5587311 GENERAL INFORMATION:
                                                                                                                                                                                                                    Sequence 34, Application US/09235451 GENERAL INFORMATION:
                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Foley & Lardner O K Street, N.W
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            28.2%;
                                                          28.2%;
43.2%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         <u>ں</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                                                        Similarity 43.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MOORE, Joan
JOHNSON, LA
WILLADSEN,
                protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3 KLFKEVKLEELENOKG-
                                                                                                                  KLFKEVKLEELENQKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Washington,
    linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 COBON,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3000 K
; TOPOLOGY: 1:; MOLECULE TYPE:
US-09-197-636-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SULT 6
-08-325-071-61
                                                                                                                                                                                          RESULT 5
US-09-235-451-34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ORGANISM: HO-09-235-451-34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ID NO 34
ENGTH: 839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      COUNTRY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Query Match
Best Local S
Matches 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRT
                                                          Query Match
                                                                          Best Local
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              319
                                                                                                                   ო
                                                                                                                                               319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT
                                                                                                                                               qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                   ð
```

```
4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 549;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICANT: JOHNSON, Law Anthony York
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 6235283man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 51; DB 1;
Pred. No. 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 29,768
R: 60042/111
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               388
                                                                                                                                                                                                                                                                                             PCT/AU87/00401
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     36
SOFTWARE: Patentin Release #1.0, CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
                                                                                                    APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                    07/926,368
                                                                                                                                                                                                                         APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/004(
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Sequence 61, Application US/08451004A
Patent No. 6235283
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 13 LENOKGSGLGKAQCAAL----WLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER: 6004
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEFAX: 904136
INFORMATION FOR SEQ ID NO: 61:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ;
9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3000 K Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Lardner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26.4%;
ilarity 35.7%;
Conservative
                                                                                                                                                                 APPLICATION NUMBER: US 0'
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE CHARACTERISTICS:
LENGTH: 549 amino acids
TYPE: amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   NAME: BENT, Stephen A REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & L
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                  PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STREET: 3000 K Štr
CITY: Washington,
COUNTRY: USA
```

ï

4

37.rai

```
Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOOKE,
JOHNSON, La.
WILLADSEN, Peter
KEMP, David Harold
SRISKANTHA, Alagacone
I: RIDING, George Alfred
RAND, Keith No. 6235283man
TON: DNA Encoding A Cell Membrane
                                                                                   COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29,768
R: 60042/111 BIAU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 51; DB Pred. No. 52;
                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: 07/242,196
FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PI4912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      412 LKNDEAAYKGONKCVKVDNLFWFQCADG 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT 9
US-08-461-004A-65
; Sequence 65, Application US/08461004A
; Patent No. 6235283
; GENERAL INFORMATION:
; APPLICANT: COBON, Stewart Gary
; APPLICANT: MOORE, Joanna Terry
                                                                                                                                                                                                                                                                        PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/242,196
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13 LENOKGSGLGKAQCAAL ---- WLQCASG 36
                                                                                                                                                                                                                                            us 08/062,109
                                                                                                                                                                                MBER: US/08/325,071
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER: AU PH9196 FILING DATE: 27-NOV-1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      REFERENCE/DOCKET NUMBER: 60(TELECOMMUNICATION INFORMATION: 202 672 5300TELEPHONE: 202 672 5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     6
  y & Lardner
Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.4%;
larity 35.7%;
Conservative
                                                                                                                                                                                                                                        APPLICATION NUMBER: US 0
FILING DATE: 17-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 27-NOV-1986 ATTORNEY/AGENT INFORMATION: NAME: BENT, Stephen A. REGISTRATION NUMBER: 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               620 amino acids
ino acid
linear
                                                                                                                                                                        CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 14-OCT-19
 ADDRESSEE: Foley & Lan
STREET: 3000 K Street,
CITY: Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEX: 904136
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   protein
                                                                                                                                                                                                                          PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               APPLICANT: RAND, KO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local Similarity
Matches 10; Conser
                                                                 20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                ; TYPE: amino & TOPOLOGY: lin; MOLECULE TYPE: US-08-325-071-65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               LENGTH:
                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         QQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Length 549;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               US-08-325-071-65

US-08-325-071-65

Sequence 65, Application US/08325071

Patent No. 5587311

GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
APPLICANT: MOORE, Joanna Terry
APPLICANT: WILLADSEN, Peter
APPLICANT: WILLADSEN, Peter
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: RAND, Reith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71

CORRESPONDENCE ADDRESS:
TYPE: Floppy disk
SR: IBM PC compatible
ING SYSTEM: PC-DOS/MS-DOS
RE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 51; DB
Pred. No. 46;
6; Mismatches
                                                                                APPLICATION NUMBER: US/08/461,004A FILING DATE: 04-JUN-1995 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                    FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29,768
R: 60042/152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        388
                                                                                                                                                                                       APPLICATION DATE: 17-MAY-1955
FILING DATE: 17-MAY-1955
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      --WLQCASG 36
                                                                                                                                                                PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
                                                                                                                                          08/325,071
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       1:11: : 1::1 : | 111 | 361 LKNQEAAYKGQNKCVKVDNLFWFQCADG
                                                                                                                                                                                                                                                                                         MBER: 07/242,196
06-JUL-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
                                                                                                                                                                                                                                                                                                                                                                                           FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                   IMBER: AU PI2570
19-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 26.4%;
llarity 35.7%;
Conservative
                                                                                                                                      APPLICATION NUMBER: US 0 FILING DATE: 19-OCT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             : 549 amino acids
amino acid
                                                                     CURRENT APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13 LENQKGSGLGKAQCAAL--
                                                                                                                                                                                                                                                                                                             FILING DATE: 06-JUL-1
PRIOR APPLICATION DATA:
APPLICATION NUMBER: E
                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER:
FILING DATE: 19-JUN-
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                protein
                                                                                                                                                                                                                                                                           PRIOR APPLICATION DATA: APPLICATION NUMBER:
                        OPERATING SYSTEM:
SOFTWARE: Patent
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Query Match
Best Local Similarity
Matches 10; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         TELEX: 904136
INFORMATION FOR SEO
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TYPE.
TOPOLOGY: LLL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          202
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 US-08-461-004A-61
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            TELEFAX:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MOLECULE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      à
```

Gaps

4;

```
Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
APPLICANT: WILLADSEN, Peter
APPLICANT: KEMP, David Harold
APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
                                                                                                                                                                                                                               COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,071
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ed. No. 55;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         60042/111 BIAU
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 51;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              MBER: PCT/AU97/00401
27-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           36
                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: US 08/062,109 FILING DATE: 17-MAY-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -WLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MBER: 07/242,196
06-JUL-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Sequence 67, Application US/08325071 Patent No. 5587311 GENERAL INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: AU P12570 FILING DATE: 19-JUN-1987 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        <u>;</u>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         29,768
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER: 60
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5399
TELEFAX: 202 672 5399
                                                                                                                                                   : Foley & Lardner 3000 K Street, N.W. shington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        th 26.4%; Similarity 35.7%; 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       650 amino acids
amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TELEFAX: 202 v.r.
TELEX: 904136
INFORMATION FOR SEQ ID NO:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           13 LENQKGSGLGKAQCAAL---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        NAME: BENT, Stephen A. REGISTRATION NUMBER: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: (FILING DATE: 06-JUL-)
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     FILING DATE: 27-NOV-1
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                  NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                    Washington,
                                                                                                                                                                                                   COUNTRY: USA
ZIP: 20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TYPE:
                                                                                                                                                     ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 11
US-08-325-071-67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        US-08-325-071-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY
                                                                                                                                                                 STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
Best Local S
Matches 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MOLECULE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 620;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                               COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       8;
Glycoprotein Of A Tick Gut 71
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ed. No. 52;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 51;
                                                                                                                                                                                                                         FILING DATE: 04-JUN-1995
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         COBON, Stewart Gary
MOORE, Joanna Terry
JOHNSON, Law Anthony York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |:||: | : | : | 412 LKNQEAAYKGQNKCVKVDNLFWFQCADG 439
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       60042/152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MBER: PCT/AU87/00401
27-NOV-1987
                                                                                                                                                                                                                                                                                    APPLICATION DATA:
FILING DATE: 19-OCT-100
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
APPLICATION 17-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 07/926,368 FILING DATE: 07-AUG-1992
                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 08/325,071 FILING DATE: 19-0CT-1994
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          -WLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                  JMBER: 07/242,196
06-JUL-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Sequence 63, Application US/08325071
Patent No. 5587311
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MBER: AU PI4912
16-0CT-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AU PI2570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        29,768
                                                                 Street, N.W
on, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         65:
                                                  Foley & Lardner
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity 35.7%;
10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     19-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     620 amino acids amino acids Y: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                 OPERATING SYSTEM: PC-D
SOFTWARE: PatentIn Rel
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N FOR SEQ ID NO:
CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER:
FILING DATE: 06-JUL-PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER:
FILING DATE: 27-NOV-
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: IFILING DATE: 16-OCT-IOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 07-AUG-:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          13 LENQKGSGLGKAQCAAL-
                NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     APPLICATION NUMBER:
FILING DATE: 19-JU
                                                                                  Washington,
  INVENTION:
                                                                                                               20007-5109
                                                               3000 K
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEX: 904136
INFORMATION FOR SE(
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TOPOLOGY: 11
MOLECULE TYPE:
                                                   ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-461-004A-65
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT 10
US-08-325-071-63
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TELEFAX:
                                                                                                  COUNTRY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       LENGTH:
                                                              STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
   Q.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                    PRIOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
```

ä

d

ð

9

```
4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                         MOORE,

MOORE,

SOHNSON, Law

KEMP, David Harold

SRISKANTHA, Alagacone

I: SRISKANTHA, Alagacone

I: RIDING, George Alfred

T: RAND, Keith No. 6235283man

TINVENTION: DNA Encoding A Cell Membrane

TOWNION: Glycoprotein Of A Tick Gut

TOWNION: Glycoprotein Of A Tick Gut
                                                                                                                                                                                                                                                                                                                                                                                      YSTEM: PC-DOS/MS-DOS
PatentIn Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 51; DB 4;
Pred. No. 55;
5; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                    CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,004A
FILING DATE: 04-JUN-1995
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    29,768
R: 60042/152
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: PCT/AU87/00401
FILING DATE: 27-NOV-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062 100
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER: US 07/926,368 FILING DATE: 07-AUG-1992
Sequence 63, Application US/08461004A
Patent No. 6235283
GENERAL INFORMATION:
APPLICANT: COBON, Stewart Gary
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        -- WLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MBER: 07/242,196
06-JUL-1988
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        AU PI2570
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   MBER: AU PH9196
27-NOV-1986
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       PI4912
                                                 COBON, Stewart Gary
MOORE, Joanna Terry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9
                                                                                                                                                                                                                                                                                                                                                                      compatible
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5390
TELEFAX: 202 672 5399
                                                                                                                                                                                                                              CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.4%;
illarity 35.7%;
Conservative
                                                                                                                                                                                                                                                                                                                              COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compati
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FILING DATE: 17-MAY-1993 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 27-NOV-1986 ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    APPLICATION NUMBER: AU PFILING DATE: 16-OCT-1987 PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      19-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        650 amino acids amino acids
                                                                                                                                                                                                                                                                 STREET: 3000 n crry: Washington, D.C.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     NAME: BENT, Stephen A REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS
LENGTH: 650 amino acid
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13 LENGKGSGLGKAQCAAL--
                                                                                                                                                                                                                                                                                                                                                                                                                                                      FILING DATE: 04-JUN-
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       APPLICATION NUMBER:
FILING DATE: 06-JUL-:
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      APPLICATION NUMBER:
                                                                                                                                                                                                                     NUMBER OF SEQUENCES:
                                                                                                                                                                                   TITLE OF INVENTION: TITLE OF INVENTION:
                                                                                                                                                                                                                                                                                                                                                                                       SYSTEM:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      linear
                                                                                                                                                                                                                                                                                                    COUNTRY: USA
ZIP: 20007-5109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      904136
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    BENT,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ; MOLECULE TYPE:
US-08-461-004A-63
                                                                                                                                                                                                                                                                                                                                                                      COMPUTER:
OPERATING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 TOPOLOGY:
                                                                                                                  APPLICANT:
APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                      SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       10;
                                                                  APPLICANT:
                                                                                                    APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 650;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                              APPLICANT: SRISKANTHA, Alagacone
APPLICANT: RIDING, George Alfred
APPLICANT: RAND, Keith No. 5587311man
TITLE OF INVENTION: DNA Encoding A Cell Membrane
TITLE OF INVENTION: Glycoprotein Of A Tick Gut
NUMBER OF SEQUENCES: 71
                                                                                                                                                                                                                                                                                                                 IBM PC compatible
SYSTEM: PC-DOS/MS-DOS
Patentin Release #1.0, Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      .;
8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ore 51; DB 1;
ed. No. 55;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REFERENCE/DOCKET NUMBER: 60042/111 BIAU TELECOMMUNICATION INFORMATION: TELEPHONE: 202 672 5399
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.4%; Score 51;
larity 35.7%; Pred. No.
Conservative 6; Mismatc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Law Anthony York
                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 08/062,109
FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      13 LENQKGSGLGKAQCAAL----WLQCASG 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      us 07/926,368
                                                                                                                                                                                                                                                                                                                                                               CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/325,07
FILING DATE: 14-OCT-1993
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FILING DATE: 06-JUL-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/0/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER: AU P14912
FILING DATE: 16-OCT-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU P12570
FILING DATE: 19-JUN-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PH9196
FILING DATE: 27-NOV-1986
                  Joanna Terry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    9
                                                                   KEMP, David Harold
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       29,768
                                                                                                                                                                                                Foley & Lardner 3000 K Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      67:
                                                                                                                                                                                                                                               COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 APPLICATION NUMBER: US 0'
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ATTORNEY/AGENT INFORMATION:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     650 amino acids amino acids
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       NAME: BENT, Stephen A REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TELEX: 904136
INFORMATION FOR SEQ ID NO: SEQUENCE CHARACTERISTICS
                                                                                                                                                                                                                                    Washington, D.C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         protein
                                JOHNSON, LA
                                                                                                                                                                                    CORRESPONDENCE ADDRESS
                                                                                                                                                                  NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                                             COMPUTER: IBM PC
COMPUTER: IBM PC
OPERATING SYSTEM:
 COBON,
MOORE,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Similarity
10; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      TYPE:
                                                                                                                                                                                                  ADDRESSEE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      RESULT 12
US-08-461-004A-63
                                                                                                                                                                                                                                                                                                                                                    SOFTWARE:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TOPOLOGY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -071-67
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     LENGTH:
                                                 APPLICANT
                                                                                                                                                                                                                       STREET:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        MOLECULE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      US-08-325
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      g
```

1;

18-09

```
CLL..
COUNTRY: ...
COUNTRY: ...
ZIP: 20007-5109
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
~~nMPUTER: IBM PC compatible
~~nMPUTER: IBM PC compatible
~~nMPUTER: TR PC-DOS/MS-DOS
~~TN Release #1.0, Version #1.25
                                                                                                                                                                                                                KEMP, David Harold
SRISKANTHA, Alagacone
RIDING, George Alfred
RAND, Keith No. 6235283man
FENTION: DNA Encoding A Cell Membrane
FENTION: Glycoprotein Of A Tick Gut
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       FILING DATE: 27-NOV-1987
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU87/00401
                                                                                                                                                                                    York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            29,768
R: 60042/152
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/325,071
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/062,109
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         INBER: US 08/062,109
17-MAY-1993
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 17-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,368
FILING DATE: 07-AUG-1992
PRIOR APPLICATION DATA:
                                                                                                7, Application US/08461004A 6235283
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 UMBER: 07/242,196
06-JUL-1988
                                                                                                                                                                                JOHNSON, Law Anthony
WILLADSEN, Peter
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 FILING DATE: 19-JUN-1987
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: AU PH9196
FILLING DATE: 27-NOV-1986
ATTORNEY/AGENT INFORMATION:
                                                                                                                                                  Stewart Gary
Joanna Terry
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5399
TELEFAX: 202 672 5399
                                                                                                                                                                                                                                                                                                                                    CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    67:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     16-OCT-1987
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                650 amino acids
amino acid
Y: linear
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                REFERENCE/DOCKET NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TELEX: 904136
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            NAME: BENT, Stephen A REGISTRATION NUMBER:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    protein
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FILING DATE: 06-JUL-PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         FILING DATE: 16-OCT-PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   APPLICATION NUMBER:
                                                                                                                                                                                                                                                                                                                        NUMBER OF SEQUENCES:
                                                                                                                                                                                                                                                                                     TITLE OF INVENTION: TITLE OF INVENTION:
                                                                                             Sequence 67, Applicat
Patent No. 6235283
GENERAL INFORMATION:
                                                                                                                                                   COBON,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                TOPOLOGY: 1:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    US-08-461-004A-67
                                                                              US-08-461-004A-67
                                                                                                                                                                                                                                  APPLICANT:
APPLICANT:
APPLICANT:
                                                                                                                                                                                   APPLICANT: APPLICANT:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  LENGTH:
                                                                                                                                                  APPLICANT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     MOLECULE
                                                               RESULT
             d
```

```
ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                               4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length 259;
  Length 650;
                                                                                                                                                                                                                                         Campylobacter Jejuni Antigens, And
Methods For Their Production And Use
11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Version #1.25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   9;
                                ъ,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       7
 4 ;
                                                                                                                                                                                                                                                                                                              SOFFEN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      DB
38;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 49; DB
Pred. No. 38;
5; Mismatches
Score 51; DB
Pred. No. 55;
6; Mismatches
  DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               NAME: Gray III, William O.
REGISTRATION NUMBER: 30,944
REFERENCE/DOCKET NUMBER: P/1261-13
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 382-0888
TELEX: 236925
INFORMATION FOR SEQ ID NO: 8:
                                                                                                                                                                                                                                                                                                            3: OSTROLENK, FABER, GERB & 1180 Avenue of the Americas
                                                                                                                                                                                                                                                                                                                                                                     COUNTRY: USA

ZIP: 10036-8403

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Very
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/402,804
FILING DATE:
CLASSIFICATION DATA:
APPLICATION NUMBER: US 08/114,420
FILING DATE: 30-AUG-1993
APPLICATION NUMBER: US 08/112,387
FILING DATE: 27-AUG-1993
                                                                                  |:||: |: |: |: 42 IKNQEAAYKGQNKCVKVDNLFWFQCADG 469
                                                             36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        APPLICATION NUMBER: US 08/112,387
FILING DATE: 27-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/986,928
FILING DATE: 08-DEC-1992
PRIOR APPLICATION DATA:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  APPLICATION NUMBER: US 07/612,330 FILING DATE: 13-NOV-1990 ATTORNEY/AGENT INFORMATION:
                                                               -WLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      27
                                                                                                                                                                    Sequence 8, Application US/08402804 Patent No. 5874300
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                             9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LFKEVKLEELENQKGSGLGKAQCA
                                                                                                                                                                                                                    Martin J
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Similarity 41.7%;
.0; Conservative
Query Match
Best Local Similarity 35.7%;
Matches 10; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          259 amino acids amino acids
                                                                                                                                                                                                 GENERAL INFORMATION:
APPLICANT: Blaser, Martin
APPLICANT: Pei, Zhiheng
TITLE OF INVENTION: Campy
TITLE OF INVENTION: Metho
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: OSTROLENK, F
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CHARACTERISTICS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          protein
                                                                13 LENOKGSGLGKAQCAAL-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           linear
                                                                                                                                                                                                                                                                                                                                           New York
New York
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TOPOLOGY: 11

MOLECULE TYPE:
US-08-402-804-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   10;
                                                                                                                                                         US-08-402-804-8
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LENGTH:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           TYPE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  134
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    4
                                                                                                                                           RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δŽ
                                                                                             QQ
                                                                 QY
```

15

RESULT

 $\infty$ 

```
US-08-470-260-3

US-08-470-260-3

US-94Gerea 2. Application US/08470260

SAUGHER N. 6077706

SAUGHER OF HUNBERTON:

SAUGHER SAUGHER SAUGHER

CATTY: Beneryll SAUGHER

CATTY: Beneryll SAUGHER

SAUGHER SAUGHER SAUGHER

CANASITE ABOUNDER:

SAUGHER SAUGHER SAUGHER

SAUGHER SAUGHER SAUGHER

CANASITE ABOUNDER:

SAUGHER SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUGHER SAUGHER

SAUG
```

.

```
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd
```

OM protein - protein search, using sw model

Seconds 2002, 09:26:16 October 24, Run on:

updates/sec ; Search time 7.0219 (without alignments) 506.317 Million cell

US-09-894-030-3\_COPY\_1\_37 193 Title:

1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG Perfect score: Sequence:

Scoring table:

37

283138 seqs, 96089334 residues BLOSUM62 Gapop 10.0 , Gapext 0.5 Searched: 283138 of hits satisfying chosen parameters: Total number

Minimum DB seq length: 0 Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

Database :

PIR\_71:\*
1: pir1:\*
2: pir2:\*
3: pir3:\*
4: pir4:\* 122...4

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

	Description	blancin 168	$\mathbf{H}$	stathionine	netic	cept	ole minc	probable tail prot	mine	C	conserved hypothet		hypothetical prote	_	ical	anscri	cal prot	etRNA 1	Ξ	ical prot	ш	casp homolog - fis	etical p	in	gglutinin	gglutinin	gglutinin	glutinin pr	de-aspartate	hypothetical prote
COTTAGENIOC	ID	T1278	4384	6142	0	<b>c</b> 762	1363	33	122	δ	025	IVF	$\sim$	2415	10	10	_		T50308	845	5429	41	00	HMIVF5	HMIVF7		٧F	A46339	¥B.	C86314
	Length DB	56	09	64	02	39	00	900	69		95	64	024	84	909	61	33	38	20	38	91	33	48	64	64	64	64	568 1	24	64
æ	Query Match	0.	•	_:	_:	_:	٠.							٠.	ö	·	Ö	'n	o.	S.	ю.	o.	ſυ.	ъ.	رى	'n.	'n	25.6	ഗ	Ś
	Score	193	61	58	26	54.5	വ	53	52	52	H.	ij.	<del>, - i</del>	51.5	Ή.	51	51	51	51	51	Ö	50.5	uı	9	9	φ.	o,	49.5	σ.	o,
	••		7	æ	4	ഗ	9	7	80	5	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	58

Length 56;

Score 193; DB 2; Pred. No. 3.4e-18;

100.0%; 100.0%;

Query Match Best Local Similarity

24	cal p ABC-t	pdhC	14 C	יי כא		a,	prot	0	lation prote	nethyltran	al	g O	vacuolating cytoto
F25	84 51	10.0	7	8	10	34	T48130	S66820	PQ0041	T37754	T28130	B53739	G64630
77	0 0	CI (	~	~	~	~	~	-	7	7	~	7	7
8 4	243 259	6	407	28	2	111	200	534	260	802	1122	28	1290
ō. <del>4</del> .	4.4	4.	₹.	4.	4.	ᆫ.	٦.	٦.	۲.	ᇅ	۲.	٦.	٦.
25	10.10	10	io i	D.	LO.	Ы	Ω	S	S	S	10	S	S
49.5	4 4 9	49	49	49	49	8	8	8	48.5	ω.	ω,	8	œ
30 31	32 33	34	35	36	37	38	39	40	41	42	43	44	45

### ALIGNMENTS

_	RESULT 1
	sublancin 168 precursor - Bacillus subtilis phage SPBc2
	C;Species: Bacillus subtilis phage SrBcz C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
	C; Accession: T12/83; H09/19 R; Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.
	submitted to the EMBL Data Library, August 1997 A; Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 pro
	A;Status: translated from GB/EMBL/DDBJ
	A; Molecule Lype: Dina A; Residues: 1-56 <laz></laz>
	A; Cross-references: EMBL: AF 020/13; NID: 93023476; FID: 9302317; Innoverse R; Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
	C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
	ontita M · Endita V · Enma
	A; Authors: Fourger, D.; Filez, C.; Fujiea, M.; Fujiea, I.; Fuma, D.; Hullo, M.; Holsappel, S.; Hosono, S.; Hullo, M.
	Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
	A; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, A.; Lasuda, S., Adda, J., Logical S., Andrews, V.; Pohl, T.M.; Portete
	Y.; Sato, T
	A; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
	mamoto, H.; Yamane, K.; Yasumoto, F
	mstein,
	equence of the Gram-Positive Dacterium Bactitus ID:98044033
	A; Status: nucleic acid sequence not shown; translation not shown
_	
	A; Cross-references: GB: Z99115; GB: AL009126; NID: 92634478; PIDN: CAB14066.1; PID: e11835
	A; Experimental source: strain 168
	A; Gene: yolg
	C; Genetics: <ku1></ku1>
	A; Gene: sunA
	Keywords: antibiotic; lanthionine
	F;1-19/Domain: propeptide #status predicted <pro> =:30-56/broduct: sublancin 168 #status predicted <mat></mat></pro>
	26-55/Disulfide bonds: #status experimental
	cus predicted anine (Ser) #status experimental
	)-3-methyl-]

Fujii, C.Y L.; Tallon,

L.; Tallon, C.M.; Venter

```
Capsaich receptor, VR1 - human
Cispecies: Homo saplens (man)
Cispecies: Jo-Jun-2001 #sequence_revision 30-Jun-2001
Ciscosion: JG7621
Richard Biophys: Res. Commun. 281, 1183-1189, 2001
A; Title: The tissue distribution and functional characterization of human VR1.
A; Contents: Dorsal root ganglia
A; Contents: Dorsal root ganglia
A; Molecule type: mRNA
A; Residues: 1-839 <COR>
A; Coress references: GB: ARI96175
C; Comment: This receptor, a transmembrane protein with many phosphorylation sites, is its integrative activation by several noxious stimuli, and plays an important role in C; Genetics:
C; Comment: This receptor, a transmembrane predicted <ANKI>
C; Comment: This ransmembrane #status predicted <ANKI>
F; 201-233, Domain: ankyrin #status predicted <ANKI>
F; 201-233, Domain: transmembrane #status predicted <ANKI>
F; 204-233, Domain: transmembrane #status predicted <ANKI>
F; 205-201-203, Domain: transmembrane #status predicted <ANKI>
F; 205-201-203, Domain: transmembrane #status predicted <ANKI>
F; 207-203, Domain: transmembrane #status predicted <ANKI>
F; 207-203, Domain: transmembrane #status predicted <ANKI>
F; 207-204, Molecule type for loop #status predicted <ANKI>
F; 207-204, Domain: transmembrane #status predicted <ANKI>
F; 207-204, Domain: transmembran
                                                                                                                                                                                                                     hypothetical protein At2g21970 [imported] - Arabidopsis thaliana C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C; Accession: D84607
R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y R; Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, Nature 402, 761-768, 1999
A; Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A; Reference number: A84420; MUID: 20083487
A; Reference number: A84420; MUID: 20083487
A; Residues: 1-202 <STO>
A; Residues: 1-202 <STO>
A; Cross-references: GB: AE002093; NID: g4417289; PIDN: AAD20414.1; GSPDB: GN00139
C; Genetics:
A; Map position: 2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2; Length 839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length 202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ore 56; DB 2 ed. No. 7.8; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Score 54.5; Di
Pred. No. 47;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SGLGKAQCAALWLQ
                              EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Score 56; Pred. No. 7
                                                              130 LFKKLDVEGLSEAIGAGLAAMGCAAMFAWL 159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   LFKEVKLEELENQKGSGLGKAQCAAL--WL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          28.2%;
larity 43.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              29.0%;
larity 43.3%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              KLFKEVKLEELENOKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        l Similarity
13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
16; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Query Match
Best Local S
Matches 16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local
                              ď
                                                                                        296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Matches
                                                                                                                                                                              4
                                                                                        g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Qγ
                                                                                                                                                                                                                                        fu
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       40
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .; PID:g560649 cellulose-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Σ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ote: F24I3.130
uperfamily: O-succinylhomoserine (thiol)-lyase
eywords: carbon-sulfur lyase; chloroplast; methionine biosynthesis; phosphoprotein;
eywords: carbon-sulfur lyase; chloroplast) #status predicted <TNP>
-70/Domain: transit peptide (chloroplast) #status predicted <MAT>
1-464/Product: cystathionine beta-lyase #status predicted <MAT>
1-464/Product: cystathionine beta-lyase (Lys) (covalent) #status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      297/2; 336/3; 372/3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ; Weichselgartner,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       cystathionine beta-lyase
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9-May-2000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ID:9704397
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                B30669.1;
fungal ce.
y <SXY>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                thaliana
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
     ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         cystathionine beta-lyase (EC 4.4.1.8) precursor - Arabidopsis thal N.Alternate names: protein F2413.130
C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 27-Apr-1996 #sequence_revision 27-Feb-1997 #text_change 19 C; Accession: S61429; T47765
C; Material S.; Ruffet, M.L.; Douce, R. Plant Mol. Biol. 29, 875-882, 1995
A; Title: Cloning of an Arabidopsis thaliana cDNA encoding cystathin A; Reference number: S61429; MUID: 96128029
A; Molecule type: mRNA
A; Residues: 1-464 < RAV>
A; Residues: 1-464 < RAV>
A; Residues: 1-464 < RAV>
A; Reference number: Z24475
A; Reference number: Z24475
A; Residues: 1-464 < NINA
A; Residues: 1-464 
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      streptomyces endo-1,4-beta-xylanase A homology; fusicity Streptomyces endo-1,4-beta-xylanase A homology; fusicity Streptomyces endo-1,4-beta-xylanase A homology < no. fungal cellulose-binding domain homology < FCB>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    면정
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    237/3; 270/1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                #text_change
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 464
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               20;
;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  836
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      78/1; 116/3; 140/2; 171/1; 200/2;
                                                              37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  9.3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  6.8;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 61; DB Pred. No. 6.8; 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                              1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                        Mismatches
     Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               801 OKVTVTVTVESQPTQGGMNQGGGNCAAKWGQCGGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         58;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        SO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              <del>"</del>
     ő
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     31.6%;
llarity 36.1%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      30.1%;
llarity 48.6%;
Conservative
     Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
Les 13; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Local Similarity
tes 17; Conserver
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   A; Accession: S43846
A; Status: preliminary
A; Molecule type: mRNA
A; Residues: 1-860 <BLA>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               278/Binding site:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Cross-references:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A; Residues: 1-860 < A; Cross-references: C; Superfamily: Stre F; 49-325/Domain: St F; 829-860/Domain: f
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               A; Introns: 21/3; 7
A; Note: F24I3.130
C; Superfamily: O-s
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
     Matches
                                                                                                                                                                                                        2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ö
                                                              δ
```

is in

.. H

g

ô

```
Conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Species: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C;Accession: D90252
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
Submitted to GenBank, April 2001
A;Deference number: April 2001
A;Deference number: Appil 39
A;Reference number: Appil 30
A;Reference number: Appil 30
A;Reference number: Appil 30
A;Refere
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         R.; De Clerck,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                385/3; 398/2; 450/3; 484/3;
A;Residues: 1-469 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99243.1; PID:g16410581; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo1165
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     24-Nov-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      hypothetical protein F26P21.30 - Arabidopsis thaliana C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Nov C;Accession: T05299
R;Bevan, M.; Terryn, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; Sewes, H.W.; Mayer, K.F.X.; Schueller, C.
submitted to the Protein Sequence Database, October 1998
A;Reference number: Z15407
A;Reference number: Z15407
A;Residues: 1-761 < BEV>
A;Residues: 1-761 < BEV>
A;Cross-references: EMBL:AL031804
A;Experimental source: cultivar Columbia; BAC clone F26P21
C;Genetics:
A;Map position: 4
A;Introns: 123/2; 167/2; 221/3; 262/3; 281/3; 348/3; 385/3; 398/2; 450
C;Superfamily: Arabidopsis thaliana hypothetical protein F26P21.30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ö
                                                                                                                                                                                                                                                                                          ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length 295;
                                                                                                                                                                                                                         469;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 761;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Indels
                                                                                                                                                                                                                                                                                            Indels
                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             4 ;
                                                                                                                                                                                                                                                                                            4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        .;
7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ñ
                                                                                                                                                                                                                         3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DB
                                                                                                                                                                                                                       ore 52; DB ced. No. 57; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 52; DB Pred. No. 90; 4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.7%; Score 51.5; I 54.2%; Pred. No. 43; ive 6; Mismatches
                                                                                                                                                                                                                         Score 52;
Pred. No. 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 293
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                KLFKEVKLEELENOKG-SGLGKAQ 25
                                                                                                                                                                                                                                                                                      3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    270 KVYEEKELEELENRLGKSSLLKAK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  26.9%;
larity 50.0%;
Conservative
                                                                                                                                                                                                                       26.9%;
61.1%;
                                                                                                                                                                                                                                                                                                                                                             MEKLFKEVKLEELENQKG 18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | : | :|:| |:|||
288 SQITKLECSAFWVQCA 303
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Conservative
                                                                                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SGLGKAQCAALWLQCA 34
                                                                                                                                                                                                                                                                                                                                                                                                 :||| |:| ||:| |||
LEKLVKKVLLEKLAEQKG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Ouery Match
Best Local Similarity
Matches 8; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local Similarity
es 13; Conser
                                                                                                                                                                                                                                                        Similarity
11; Conser
                                                                                                                                                                                                                       Query Match
Best Local S
Matches 11
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             13
                                                                                                                                                                                                                                                                                                                                                                 Ч
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Φ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 RESULT
T05299
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    D90252
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                đ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AESOLATOR AND ACCESSION TO A STATE AND A S
                                                                                                                                                                                                                                                                                                                                                                                                                                  s bacteriophage phi Sfil
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         bacteriophage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PIDN: AAC34413
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Ω
                                                                                                                                                                                                                                                                                              13-Aug-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       24-Sep-1999
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    probable tail protein - Streptococcus phage phi-01205
C; Species: Streptococcus phage phi-01205
C; Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
C; Accession: T13331
R; Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, I
Microbiology 143, 3417-3429, 1997
A; Title: Sequence analysis and characterization of phi 01205, a temperate k
A; Reference number: 217654; MUID: 98048466
A; Reference number: 217654; MUID: 98048466
A; Status: translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-1006 <STA>
A; Residues: 1-1006 <STA>
A; Residues: I-1006 <STA>
A; Experimental source: host Streptococcus thermophilus strain CNR21205
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       2;
                                                                                                                                                                                                               probable minor structural protein - Streptococcus phage phi-Sfill
C; Species: Streptococcus phage phi-Sfill
C; Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change
C; Accession: T13636
R; Lucchini, S.; Desiere, F.; Bruessow, H.
Virology 246, 63-73, 1998
A; Title: The structural gene module in Streptococcus thermophilu
A; Reference number: Z17696; MUID: 98321150
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  A; Residues: 1-1000 <LUC>
A; Cross-references: EMBL: AF057033; NID: 93320432; PID: 93540286; Experimental source: specific_host Streptococcus thermophilus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 1000
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 1006
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     13;
                                                 355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-1000 < LUC>
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 53; DB Pred. No. 86; 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  ore 53; DB ed. No. 87; Mismatches
                   499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                470 LNEMQGLKGRRVGNGQCYALAAWYSMKLGG 499
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        10 LEELENQKGSGLGKAQCAAL - WLQCASGG 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            10 LEELENQKGSGLGKAQCAAL--WLQCASGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  27.5%;
llarity 40.0%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   / Match
Local Similarity 40.0%;
nes 12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Local Similarity
nes 12; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Best Loc
Matches
                                                 319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              œ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
```

g

ğ

g

ð

492/

ô

ij

ж ::

2002

14:40:04

24

Oct

Thu

```
hypothetical protein R11A8.7b - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Accession: T24158; T24179
C; Accession: T24158; T24179
R; Bardill, S.
Submitted to the EMBL Data Library, April 1996
A; Reference number: 219846
A; Reference number: 219846
A; Reference number: 219846
A; Recension: T24158
A; Scatus: preliminary; translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 1-2584 <WIL>
A; Cross-references: EMBL: 270686; PIDN: CAB54289.1; GSPDB: GN00022; CESP:R11A8.7b
A; Cross-reference: clone R10H10
R; Cummings, P.
Submitted to the EMBL Data Library, March 1996
A; Reference number: 219849
A; Accession: T24179
A; Accession: T24179
A; Accession: T24179
A; Residues: 1-2584 <WIZ>
A; Reperimental source: clone R11A8
A; Residues: 1-2584 <WIZ>
A; Reperimental source: clone R11A8
A; Reperimental source: clo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                hypothetical protein R11A8.7a - Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; Species: Caenorhabditis elegans
C; C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C; Accession: T24157; T24177
R; Bardill, S.
Submitted to the EMBL Data Library, April 1996
A; Reference number: 219846
A; Reference number: 219846
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Rossidues: 1-2606 < WILL>
A; Cross-references: EMBL: 270686; PIDN: CAA94615.1; GSPDB: GN00022; CESP: R11A8.7a
A; Experimental source: clone R10H10
B; Cummings, P.
Submitted to the EMBL Data Library, March 1996
A; Reference number: 219849
A; Reference number: 219849
A; Reference number: 21060 < WILL>
A; Relatus: preliminary; translated from GB/EMBL/DDBJ
A; Status: preliminary; translated from GB/EMBL/DDBJ
A; Rosidues: 1-2606 < WILL>
A; Residues: 1-2606 < WILL>
A; Residues: 1-2606 < WILL>
A; Reperimental source: clone R11A8
A; Reperimental source: clone R11A8
A; Rapen: CESP: R11A8.7a
A; Map position: 4
A; Map position: 4
A; Introns: 24/3; 80/3; 137/2; 177/1; 218/3; 348/3; 507/2; 709/3; 753/2; 924/1; 1020
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length 2584;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length 2606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Score 51.5; DB 2;
Pred. No. 3.3e+02;
3; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 51.5; DB 2;
Pred. No. 3.3e+02;
3; Mismatches 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                12 ELENQKGSGLGKAQCAALWLQCASG 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12 ELENQKGSGLGKAQCAALWLQCASG 36
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           26.7%; ; ; iilarity 40.0%; i Conservative 3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       26.7%;
larity 40.0%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity
nes 10; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 10; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  :: |:||
1485 DVRNKKG--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            RESULT 14
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           QΩ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  qq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ð
                                                                                                                                                                                                                                                                                                                                                                                                           enes in different region
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                l, S.; White, O.; Alonso Creasy, T.H.; Dewar, K.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   S.; Khaykin, E.; Kim, C
S.; Maiti, R.; Marziali
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      H.; Tallon,
                                                                                                                                   bama/1/75 [H4N8])
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 pouq
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ï
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GSPDB:GN00141
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 thiolester
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      A.M.; Sun,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     11-Dec-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            status predicted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   dopsis.
                                                                                                                                                     N:Contains: hemagglutinin HA1; hemagglutinin HA2
C; Species: influenza A virus
C; Species: influenza A virus
C; Species: influenza A virus
C; Species: liftuenza A virus
C; Species: liftuenza A virus
C; Species: liftuenza A virus
C; Accession: F34214
R; Donis, R.O.; Bean, W.J.; Kawaoka, Y.; Webster, R.G.
Virology 169, 408-417, 1989
A; Title: Distinct lineages of influenza virus H4 hemagglutinin genes
A; Reference number: A34214; MUID:89204912
A; Reference number: A34214; MUID:89204912
A; Residues: 1-564 < DON>
A; Residues: 1-564 < DON>
A; Residues: 1-564 < DON>
A; Cross-references: GB:MZ5288; NID:g324228; PIDN:AAA43221.1; PID:g324
C; Genetics:
A; Map position: segment 4
C; Superfemily: influenza virus hemagglutinin
C; Keywords: glycoprotein; hemagglutinin HA1 #status predicted < HA1>
F; 11-16/Domain: signal sequence #status predicted < HA2>
F; 344-564/Product: hemagglutinin HA2 #status predicted < HA2>
F; 344-564/Product: hemagglutinin HA2 #status predicted < HA2>
F; 344-564/Product: hemagglutinin #A2 #status predicted < HA2>
F; 345-564/Fidinding site: carbohydrate (Asn) (covalent) #status F; 26-480, 64-291, 76-88, 151-487, 295-319/Disulfide bonds: #status predicted F; 553, 560, 563/Binding site: palmitate (Cys) (covalent) #status predicted F; 553, 560, 563/Binding site: palmitate (Cys) (covalent) #status predicted Cys)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         hypothetical protein F6F9.12 - Arabidopsis thaliana C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Species: Arabidopsis thaliana (mouse-ear cress)
C; Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-De C; Accession: E86331
R; Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; R; Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creas ansen, N.F.; Hughes, B.; Hulzar, L.; Conway, A.B.; Conway, A.R.; Kaulere 408, 816-820, 2000
A; Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; K., Liu, Z.A.; Liu, Z.A.; Luros, J.S.; M. Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M. A; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M. A; Reference number: A86141; MUID:21016719
A; Reference number: A86141; MUID:21016719
A; Residues: preliminary
A; Molecule type: DNA
A; Residues: 1-1024 <STO>
A; Cross-references: GB:AE005172; NID:g10086489; PIDN:AAG12549.1; GSPD C; Genetics:
A; Map position: 1
                                                                                                                   hemagglutinin precursor - influenza A virus (strain A/chicken/Ala N; Contains: hemagglutinin HA1; hemagglutinin HA2
C; Species: influenza A virus
C; Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 1
C; Accession: F34214
R; Donis, R.O.; Bean, W.J.; Kawaoka, Y.; Webster, R.G.
Virology 169, 408-417, 1989
A; Title: Distinct lineages of influenza virus H4 hemagglutinin ge
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Length 564;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length 1024
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB 2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ed. No. 1.4 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 51.5;
Pred. No. 79;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             789
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Score 51.5;
Pred. No. 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |:|||: |: || : || 458 MDKLFERVRRQLRENAEDKGNG---CFEIFHQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EKLFKEVKLEELENQKGSGLGKAQCAALWLQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           763 EKLLAEVKSDLESAOKSNGMGETO--
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         26.7%;
36.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    26.7%;
ilarity 40.6%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1 Similarity
12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Similarity 13; Conser
```

Query Match Best Local

Local

Matches

g

ö

924/1; 1020/3

Query Match Best Local

RESULT

d

ð

```
RESULT 15
C83350
probable transcription regulator PA2359 [imported] - Pseudomonas aeruginosa (strain PA01
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 23-Mar-2001
C;Accession: C83350
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Linn,
J. Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathch; A;Accession: C83350
A;Accession: C83350
A;Status: preliminary
A;Molecule type: DNA
A;Mosicule type: DNA
A;Cross references: GB:AE004662; GB:AE004091; NID:g9948395; PIDN:AAG05747.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2359
C;Superfamily: Pseudomonas syringae hrpS protein; RNA polymerase sigma factor interactic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ó
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ô
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Query Match

26.4%; Score 51; DB 2; Length 361;
Best Local Similarity 36.7%; Pred. No. 60;
Matches 11; Conservative 6; Mismatches 13; Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Search completed: October 24, 2002, 09:29:48
Job time : 11.0219 secs
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     8 VKLEELENQKGSGLGKAQCAALWLQCASGG 37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             õ
```

		ė			7.5
•	•				
			,		
			•		
				(÷	
					•
					•

```
GenCore version 5.1.3 (c) 1993 - 2002 Compugen Ltd.
                   Copyright
```

OM protein - protein search, using sw model

Seconds October 24, 2002, 09:24:01; Search time 4.32117 (without alignments) 331.536 Million cell u Run on:

updates/sec

US-09-894-030-3\_COPY\_1\_37 193 1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG Title: Perfect score:

Sedneuce:

Scoring table:

Searched:

105224 seqs, 38719550 residues 0.5 BLOSUM62 Gapop 10.0 , Gapext

105224 Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries

SwissProt\_40:\* Database:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

													•																				
Descript	53780 arabidopsi	13311 bacteriop	nft7 drosophi	19695 infl	583 aquifex	Lv3 homo sap	c12 hom	37	143 influenza	7007	701	2	136	156 bos t	678 campy	746 homo sapi	5 helicobac	247 helicobac	939 bacillus	1352 drosophi	P05782 xenopus lae	383 s	7759 azorhiz	42832 sc	8245 hel	5981	8258 helicoba	8415 arc	939 lact	5918 homo sap	5466 homo sapie	60 caenorha	571
SUMMA	, ,	YOGE_BPT4	HXK2_DROME	HEMA_IACKA	Œ	CI21_HUMAN	- 1	HEMA_IAMAA	- 1	- 1	HEMA_IASE2	•	HEMA_IAMAB	ASPH_BOVIN	PEB1_CAMJE	CMC1_HUMAN	- 1	HEL	O	DJ_DROME	K1C3_XENLA	NAB4_YEAST	- 1	ω,	VAC2_HELPY	VACA_HELPY	VAC4_HELPY	YI64_ARCFU	XYLB_LACPE	RAG1_HUMAN	SHP_HUMAN	YNP7_CAEEL	ERR2_HUMAN
. 80	i	-		<b>~</b>	Ч	-	-	٦	-	-	Н	Н	<del>,                                    </del>	-	٦	1	<del>, - 1</del>	<b>,</b>	-	٦	Н	Н	⊣	<b>-</b> -i	<b>~</b>	-	Н	<del>-  </del>	<del>, - 1</del>	Н	<del>,-4</del>	٦	1
ength	ı w	181	$\blacksquare$	10	638	$\mathbf{G}$	in	563	Ō	564	564	9	9	S	S	$\sim$	1288	29	$\overline{}$	4	280	534	260	0	28	1290	29	0	0	1043	257		
% Query Match	<u>:</u>	27.7	27.7	26.7			10	25.6	'n	٠.	'n	٠,		'n.	'n	'n.	25.4	in.	ທ	10	25.1	ъ.	2	2	کا	'n.	<u>ي</u>	₹.	4	₹.	₹.	4	₽.
Score	58	ж	ж	51.5	ស	ö	9.	49.5	о О	6	٠	σ.	σ,	φ.	4	49	4	4	ф	œ.	48.5	ω.	ω	ω.	<del>.</del>	ω.	ω.	48	48	48	47.5	47	47
sult No.		7	æ	4	ស	9	7	æ	o o	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

P17519 potato leaf P11622 potato leaf		Q64727 mus musculu p07648 escherichia				•	P26094 influenza a	P26095 influenza a
V70K_PLRV1 V70K_PLRVW		•			YRP]		HEM	HEMA_IAHC7
	55							
639 639	1065 1065	1065	1691	1938	32	47	57	57
24.4 24.4	24.4	24.4	24.4	24.4	24.1	24.1	24.1	24.1
47	47	47	47	47	9	9	46.5	9
34 35	36	800	40	41	42	43	44	45

### ALIGNMENTS

	(CBL)	ohyta; Rosidae;	ne beta-	н в., г.,	Nyakatura G., il B., siek G., Climent J., Duchemin D., D., Iberta E., Mewes HW., Jenkins J.,	cen S., V., Venter J.C., Imizu E., Y., Y.,
	4.4.1.8)	Tracheor dicots;	cystathionine coli.";	Unseld M., M., Obermaier igdomenech P. C., Brottier W., Quetier F Benes V.,	r. P., Nyak nionati B. Nordsiek Nordsiek J., Cli r. B., Duck Masuy D., Masuy D., Zasacubert ale D.,	ujii C.Y., Shea T.P. n J., Van Aken S., Feldblyum T.V., , White O., Venter J Kato T., Asamizu E., K., Kishida Y., A., Muraki A., C., Wada T.,
	date) update) precursor (EC	o o	cDNA encoding cy in Escherichia c	Columbia;  6720; PubMed=11130713;  [I., Lemcke K., Rieger M., Ansorge W., Unseld M., Valle G., Bloecker H., Perez-Alonso M., Obermaier Boutry M., Grivell L.A., Mache R., Puigdomenech P., Choisne N., Artiguenave F., Robert C., Brottier Cattolico L., Weissenbach J., Saurin W., Quetier F. Mueller-Auer S., Gabel C., Fuchs M., Benes V., Drzonek H., Erfle H., Jordan N., Bangert S.,	Angelo M., Pallavicini A., Toppo S., Simionati B., Angelo M., Pallavicini A., Toppo S., Simionati B., Iornischer K., Kauer G., Loehnert TH., Nordsiek G., Scharfe M., Schoen O., Bargues M., Terol J., Climent J. Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D., Naarse A.C., Alcaraz JP., Cottet A., Casacuberta E., Argiriou A., Flores M., Liguori R., Vitale D., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes HW., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,	Lazo M., Walts A., Utterback T., Fujii C.Y., Shea Haas B., Maiti R., Wu D., Peterson J., Van Aken Scher J., Sellers P., Gill J.E., Feldblyum T.V., In X., Nierman W.C., Salzberg S.L., White O., Ven Kaneko T., Nakamura Y., Sato S., Kato T., Asamiz Kimura T., Idesawa K., Kawashima K., Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakazaki N., Shinpo S., Takeuchi C., Wada T., Yamada M., Yasuda M., Tabata S.; Iant Arabidops Janalysis of chromosome 3 of the plant Arabidops
464 AA	update) on updat it precu	; Em] ledo: ; Ari	NA e	nsor rez- ache F., J.,	and Top ehneehne ' Ot' Pur Pur Gouor	ck Tett Pet 1 J.
4	ο	ess) hyta coty ceae		., A , Pe ., M ave ach l C.	HOII A., LO, LO Z.A. C., C., C., -P.	erba D., Gil alzba, Ka W., M., Tab
PRT;	Created) Last sequence update) Last annotation update) ase, chloroplast precurs (Cysteine lyase).	F24I3.130. thaliana (Mouse-ear cress). iridiplantae; Streptophyta; Embrya; Magnoliophyta; eudicotyledons; Brassicales; Brassicaceae; Arabi	M N.A. OLUMBIA; 8029; PubMed=8541513; Ruffet M.L., Douce R.; an Arabidopsis thaliana ctional complementation 101. 29:875-882(1995).	olumbia; 6720; PubMed=11130713; [., Lemcke K., Rieger M., Valle G., Bloecker H., Boutry M., Grivell L.A., Choisne N., Artiguenav Cattolico L., Weissenbac Mueller-Auer S., Gabel	Kranz H., Voss H., Ho Angelo M., Pallavicini A ornischer K., Kauer G., Scharfe M., Schoen O., Collado C., Perez-Perez udie M., Berger-Llauro C Maarse A.C., Alcaraz J Argiriou A., Flores M., Kaul S., Town C.D., Ko	zzo M., Walts A., Utterback T., Haas B., Maiti R., Wu D., Peters Scher J., Sellers P., Gill J.E., n X., Nierman W.C., Salzberg S.L. Kaneko T., Nakamura Y., Sato S., Kimura T., Idesawa K., Kawashima Kohara M., Matsumoto M., Matsuno Nakazaki N., Shinpo S., Takeuchi Yamada M., Yasuda M., Tabata S.; analysis of chromosome 3 of the 0-822(2000).
; Q	rea last last le,	). Mouse- tae; St lophyta les; Br	M N.A. OLUMBIA; 8029; PubMed=8541513; Ruffet M.L., Douce R. an Arabidopsis thalia ctional complementati iol. 29:875-882(1995)	Med=111 K., Ri K., Bloec Grive N., Ar L., We L., We	H., VOS K., KS K., KS M., Sch M., Pere Berger C., Alc A., Flc	waits / Maiti / Selle erman v. Naki ., Ides v. Mats v. / Mats v. / Yas v. / Sol ol).
STANDARD;	34, 34, 41, a-1y <sup>2</sup> se) (	3.13( ana ( plant gnol; sical	iA; Pubh t M.1 abido al co	lumbia; 720; PubMed=: , Lemcke K., Valle G., Blo outry M., Gr Choisne N., attolico L., Mueller-Auer	anz los M. cher rfe los C. M. cher rfe los C. do	zzo M., wal Haas B., Ma scher J., S n X., Niern Kaneko T., Kimura T., Kohara M., Nakazaki N., analysis c
ST	(Rel. 34, Cres (Rel. 34, Last (Rel. 41, Last he beta-lyase,	F24I thali firidi fa; Ma Brass	OM N.A. COLUMBIA; 28029; Pu Ruffet N an Arabi nctional Biol. 29:	Columbia; 16720; Publ M., Lemcke , Valle G., Boutry M. ., Choisne Cattolico , Mueller-	·E · E · O · ·	Rizzo M., Walts , Haas B., Mait litscher J., Sel Lin X., Nierman , Kaneko T., Nal , Kimura T., Id , Kohara M., Ma , Yamada M., Ya and analysis of
ARATH 0;	-OCT-1996 -OCT-1996 -MAR-2002 stathionin	050 OR lopsis ota; V tophyt ds II;	FR 7. 961	= CV . CV		T., T.H. T.H. D., C.M. C.M. TO S. WA C. WA
मन् मन् ।	01-OCT-1996 01-OCT-1996 01-MAR-2002 Cystathionin (Beta-cystat	AT3G57050 OR Arabidopsis t Eukaryota; Vi Spermatophyta eurosids II;	(1) SEQUENCE STRAIN=CV MEDLINE=9 Ravanel S "Cloning lyase by Plant Mol [2]	SECOENCE FRO STRAIN=CV. C MEDLINE=2101 Salanoubat M Fartmann B., Delseny M., De Simone V. Wincker P., Schaefer M.,	Wiedelmann Vezzi A., Conrad A., Reichelt J Navarro P. Cooke R., de Haan M. Monfort A.	Rooney T. Creasy T. Pai G., M Preuss D. Fraser C. Sasamoto Kiyokawa Nakayama Watanabe "Sequence "Sequence
SUI TC_			RN RC RX RT RT RU		R R R R R R R R R R R R R R R R R R R	RA RA RA RA RA RT RT

.rsp

```
EMBL; AF158101; AAD42672.1;
Hypothetical protein.
SEQUENCE 181 AA; 20681 M
                                                                                                   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE
                                                                                                                    Local
                                                                                                                                                                                                                                                                                                                     HXK2_DROME
                                                                                                                                                                                       ~
                                                                                                                                                                                                                                                                                                   RESULT
                                                                                                                                                                                        δ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a collaboration
                                                                                                                                                                                                                                                     for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       collaboration
                                                                                                                                                                                                                                                                        ib.ch/announce,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 b.ch/announce/
                                                                                                                                                                                  EMBL outstation
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL outstation trictions on it
            + NH(3)
                                                                                                                                                                                                                                 20
                                                                                                                                                                                                       (ctions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ļ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pyridoxal phosphate; Chloroplast;
                                                                                                                     FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  l J.K.;
early 1
            ine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ae:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      9;
                                                                                                                                                                                                                                is
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            This SWISS-PROT entry is copyright. It is produced through a between the Swiss Institute of Bioinformatics and the EMBI the European Bioinformatics Institute. There are no restricuse by non-profit institutions as long as its content is modified and this statement is not removed. Usage by and sentities requires a license agreement (See http://www.isb-sikor send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            dsDNA viruses, no RNA stage; Caudovirales; Myovirid
                                                                                                                                                                                                      the European contintitutions as tony wand use by non-profit institutions as tony word. Usage by and modified and this statement is not removed. Usage by and entities requires a license agreement (See http://www.isb-sientities requires to license@isb-sib.ch).
         L-homocyste
                                                                                                                                                  This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Enthe European Bioinformatics Institute. There are no restruse by non-profit institutions as long as its content
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-JAN-1990 (Rel. 13, Created)
01-JAN-1990 (Rel. 13, Last sequence update)
01-MAR-2002 (Rel. 41, Last annotation update)
Hypothetical 20.7 kDa protein in vs-regB intergenic region
106E OR VS.1 OR 61.4.
                                           COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
PATHWAY: THIRD STEP IN METHIONINE BIOSYNTHESIS.
SUBCELLULAR LOCATION: Chloroplast.
SIMILARITY: BELONGS TO THE TRANS-SULFURATION ENZYMES
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Henderson E.E., de Riel of the 58.3 to 65.5-kb
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Kutter E., Arisaka F., Kunisawa T., Tsugita A., Mosig G.
Mesyanzhinov V., Ruger W., Stidham T., Thomas E.;
"Bacteriophage T4 genome analysis.";
Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CYSTATHIONINE BETA-LYASE.
PYRIDOXAL PHOSPHAMM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (RIDOXAL PHOSPHATE (BY S DE756848549D2CA6 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10;
        + H(2)0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 58; DB 1
Pred. No. 2.5;
2; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FKLAKEVYF--LQNSEGSGLAPFDC---WL-CLRG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    181
                                                                                                                                                                                                                                                                                                                            EMBL; L40511; AAA99176.1; -.
EMBL; AL138655; CAB72175.1; -.
HSSP; P00935; 1CS1.
InterPro; IPR000277; Cys_Met_Meta_PP.
Pfam; PF01053; Cys_Met_Meta_PP; 1.
PROSITE; PS00868; CYS_MET_METAB_PP; 1.
Methionine biosynthesis; Lyase; Pyrido:
Transit peptide.
      Cystathionine
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Valerie K., Stevens J., Lynch M., Henc
"Nucleotide sequence and analysis of to
of bacteriophage T4.";
Nucleic Acids Res. 14:8637-8654(1986)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A. MEDLINE=87066735; PubMed=3024113;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               464 C
278 P
50429 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       30.18;
larity 48.68;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EMBL; X04567; CAA28228.1;
  ACTIVITY:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 278
464 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | Similarity
17; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Viruses; dsDNA vir
T4-like phages.
NCBI_TaxID=10665;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Bacteriophage
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EQUENCE FROM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  P13311;
01-JAN-1990
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                YOGE_BPT4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               TRANSIT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      296
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         YOGE_BPT4
DE CCC CCC REPRERENTE REPRESENTATION OF THE PROPERTY OF THE PR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ρŢ
```

```
ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N. A.

SEQUENCE FROM N. A.

SEQUENCE FROM N. A.

Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Adams L.D., Celniker S.E., Holt R.A., English R.A., Galle R.F.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
Amanatides P.G., Scherer S.E., Richards S., Ashburner M., Henderson S.N.,
B. George R.A., Lewis S.E., Richards S., Zhang O., Chen L.X.,
B. Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
April J.F., Agbayani A., An H.-J., Andrews Pfannkoch C., Baddwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.N.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.N.,
Ballew R.M., Carlen M.R., Bouck U., Brokstein P., Brothardy S.,
Burtis K.C., Busam D.A., Buller H., Cadieu E., Center A., Chandra I.,
Burtis K.C., Busam D.A., Buller H., Cadieu E., Center A., Chandra I.,
Burtis K.C., Busam D.A., Buller H., Cadieu E., Center A., Chandra I.,
Burtis K.C., Busam D.A., Buller G., Davenport L.B., Davies P.,
A burkon K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C.,
Burtis K.G., Earngelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
A capilos B., Dalcher A., Howland T.J., Heartis M.,
A clodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
A dalali M., Kalush F., Rarpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Balali M., Kalush F., Rarpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Burtis M.D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
A Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPhy L. Li J., Li S., Liang Y., Lin X.,
A Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
A Balazolo M., Pittman G.S., Pan S., Polly D., Weiler H.,
Belazzolo M., Pittman G.S., Pan S., Polly D., Weiler H.,
Shie B.C., Siden-Kiamos I., Simpson M., Stupski M.P., Shen H.,
Spier E., Spradling A.C., Stapleton M., Strong R.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HXK2_DROME STANDARD; PRT; 516 AA.

99NFT7; Q9NFT8; Q9VBF1;
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
16-OCT-2001
                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                                1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         four hexokinases
                                                                                                                   Length 181;
                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Duvernell D.D., Eanes W.F.; "Contrasting molecular population genetics of four hexok Drosophila melanogaster and Drosophila simulans."; Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                STRAIN-OREGON-K;
Deobagkar D.D., Kulkarni G.V., Deobagkar D.N.;
Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases
               CRC64;
                                                                                                                                                                                                                             15;
      609705684A6F99B0
                                                                                                                Ξ,
                                                                                                                                                                                                                                                                                                                                           37
                                                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                                                                                                 EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                                                                                                                                                                                                                                                                                                                                      Score 53.5;
Pred. No. 4
                                                                                                                                                                                                                       5;
20681 MW;
                                                                                                             27.78;
41.78;
                                                                                                                                                                                                                          Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STRAIN-VARIOUS STRAINS;
                                                                                                                                                               Similarity 15; Conservated
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A
```

.rsp

37

copy\_1

-030 - 3

us-09-894

3

```
1;
                                                                                                                                                                                                                                                                 collaboration
                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                              Zheng L.,
Smith H.O.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                  -phosphate
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  33_23.1,
'E3_12,
IM(S)E3_
   ng X.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ).
3 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           GLUCOSE-BINDING (POTENTIAL).
P -> A (IN STRAIN BERKELEY).
S -> N (IN STRAIN HFL97E3_15).
S -> T (IN STRAIN ZIM(S)E3_24).
S -> N (IN STRAIN SC96E3_12.3
ZIM(S)E3_35).
SKL -> NKI (IN STRAINS DPF96E3_SKL -> NKI (IN STRAINS DPF96E3_SC96E3_12.3, HFL97E3_8, HFL97E3_HFL97E3_16, ZIM(S)E3_24 AND ZIM MISSING (IN REF. 1).
      Weissenbach J.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yac Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "The genome sequence of Drosophila melanogaster."; Science 287:2185-2195(2000).

-! CATALYTIC ACTIVITY: ATP + D-hexose = ADP + D-hexose 6-F. -! PATHWAY: FIRST STEP OF SEVERAL METABOLIC PATHWAYS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       51
                                                                                                                                                                                  9
                                                                   Shao O.
Zhu X.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Indels
     Wang A.H.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MISSING (IN REF. 1
01A0492BCDDD567C
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ATP-binding
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ore 53.5; Ded. No. 11; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                    1; ALT_SEO.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; FB9110342/10; Rex Cz.
InterPro; IPR001312; Hexokinase.
Pfam; PF00349; hexokinase; 1.
PRINTS; PR00475; HEXOKINASE.
PRODOM; PD001109; Hexokinase; 1.
PROSITE; PS00378; HEXOKINASES; 1.
Transferase; Kinase; Glycolysis; AT Multigene family; Polymorphism.
BINDING 158 ATP.
DOMAIN 197 223 GLUCOS VARIANT 224 224 S -> P VARIANT 282 282 S -> N VARIANT 410 410 S -> N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                               F257590; AAG22892.1; AL
F257591; AAG22894.1; AL
F257592; AAG22896.1; AL
F257594; AAG22898.1; AL
F257594; AAG22900.1; AL
F257596; AAG22904.1; AL
F257596; AAG22906.1; AL
F257599; AAG22906.1; AL
F257599; AAG22910.1; AL
F257600; AAG22912.1; AL
F257601; AAG22916.1; AL
F257604; AAG22916.1; AL
F257606; AAG22920.1; AL
F257607; AAG22926.1; AL
F257608; AAG22928.1; AL
F257608; AAG22928.1; AL
ACC2928.1; AL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   27.7%;
illarity 44.4%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57122
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                A)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Similarity 16; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                     AF257590;
AF257591;
AF257592;
AF257593;
AF257596;
AF257596;
AF257596;
AF257590;
AF257500;
AF257600;
AF257601;
AF257602;
AF257604;
AF257606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12
516 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            FlyBase; F
InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                            EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               EMBL;
EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best
                                                                                                                THE THE THE TENT AND DAY AND D
```

CHAINS

Influenza A virus (strain A/Chicken/Alabama/1/75). Viruses; ssRNA negative-strand viruses; Orthomyxoviridae; Influenza virus A and B group; Influenza A viruses; Influenza A virus.

NCBI\_TaxID=11338;

01-FEB-1991 (Rel. 17, Created) 01-FEB-1991 (Rel. 17, Last sequence update) 16-OCT-2001 (Rel. 40, Last annotation update) Hemagglutinin precursor [Contains: Hemagglutinin HA1 Hemagglutinin HA2 chain].

A

STANDARD;

HEMA\_1A(

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                 SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

MEDLINE=89204912; PubMed=2705304;

Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.;

Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.;

"Distinct lineages of influenza virus H4 hemagglutinin genes in different regions of the world.";

Virology 169:408-417(1989).

-!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VICELL RECEPTORS AND FOR INITIATING INFECTION.

CELL RECEPTORS AND FOR INITIATING INFECTION.

(HA1 AND HA2) LINKED BY A DISULFIDE BOND.

(HA1 AND HA2) LINKED BY A DISULFIDE BOND.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Threonyl-tRNA synthetase (EC 6.1.1.3) (Threonine--tRNA ligase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      564;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        N-LINKED (GLCNAC. ..) (POT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 11;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                HEMAGGLUTININ HAI HEMAGGLUTININ HAZ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        ¥,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Score 51.5;
Pred. No. 22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       638
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            MEKLFKEVKLEELENOKGSGLGKAOCAALWLOC
                                                                                                                                                                                                                                                                                                                                                                                     EMBL; M25288; AAA43221.1; -.
PIR; F34214; HMIVF6.
HSSP; P03437; 2HMG.
InterPro; IPR001364; Hemagglutn.
Pfam; PF00509; Hemagglutinin; 1.
PRINTS; PR00329; HEMAGGLUTN12.
ProDom; PD000225; Hemagglutn; 1.
Envelope protein; Hemagglutinin; Gl
SIGNAL 1 342 HEMAGC
CHAIN 344 564 HEMAGG
CARBOHYD 18 18 N-LINK
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   178 N
310 N
497 N
63264 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Similarity 36.4%;
12; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SYT_AQUAE STAN
067583;
16-OCT-2001 (Rel. 4
16-OCT-2001 (Rel. 4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             17
344
18
34
178
310
497
564 7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    aeolicus
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       OR AQ_1667
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Aquifex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (ThrRS)
THRS OR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            OUAE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      AC
DT
DT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DE
DE
GN
OS
Π
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        q
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
```

ij

479

g

δ

514

35

MEKLFKE-VKLEELENQKGSGLGKAQCAALWLQCAS 

Oct

S., Ohtsubo M., Yoshimura A.;
a novel p21(Cip1/Waf1)-interacting

264:457-464(1999)

ciz1.

```
Mitsui K., Matsumoto A., Ohtsuka "Cloning and characterization of zinc finger protein, cizl.":
                                                                   Biophys. Res. Commun.
                                                 zinc finger protein,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Similarity 11; Conser
                                                                                                                                                                                  medulloblastoma.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          232
232
232
500
550
558
634
634
735
810
810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               276
799
741
97
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEKLFKE---
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Zinc-finger;
                                                                                                         OF.
                                                               Biochem.
                                                                                                        SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       VARSPLIC
VARSPLIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  CONFLICT
CONFLICT
CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CONFLICT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ZN_FING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Local
                                                                                                                                                                                                                                                                              Gilley
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              501
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Matches
                                                                                                                                                                                                                                                                                                                                                                                                              δy
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              d
                                                                                                                                                                                                                                                                                                                                                                   a collaboration
BL outstation -
ictions on its
                                                                                                                                                                                                                                                                                                                                                                                                                                   маγ
                                                                                                                                                                                                                                                                                                                                                                                                                                                  commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                               for commercial b.ch/announce/
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         EMBL; AE000755; AAC07549.1; -..

R EMBL; AE000755; AAC07549.1; -..

R InterPro; IPR002106; AA_tRNA_ligase_II.

R InterPro; IPR004154; HGTP_anticodon.

R InterPro; IPR004314; tRNA-synt_2b.

R InterPro; IPR002314; tRNA-synt_thr.

R Pfam; PF03129; HGTP_anticodon; 1.

R Pfam; PF03129; HGTP_anticodon; 1.

R Pfam; PF03129; HGTP_anticodon; 1.

R Pfam; PF03129; AA_TRNA_LIGASE_II_1; 1.

R PROSITE; PS00179; AA_TRNA_LIGASE_II_2; 1.

R PROSITE; PS00179; AA_TRNA_LIGASE_II_2; 1.

R PROSITE; PS00139; AA_TRNA_LIGASE_II_2; 1.

R Mainoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Metal-binding; Zinc; Complete proteome.

T DOMAIN 336 336 ZINC (CATALYTIC) (BY SIMILARITY).

T METAL 387 387 ZINC (CATALYTIC) (BY SIMILARITY).

T METAL 513 SIN ZINC; CATALYTIC) (BY SIMILARITY).

T METAL 513 SIN ZINC; CATALYTIC) (BY SIMILARITY).

T SEQUENCE 638 AA; 74097 MW; 5ECBAR227C047955 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                   ou
                                                                                                                                                                                                                                                                                                                                   E FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                               A.L.,
Huber R.
                                                                                                                                                                                                                                                                                                                                                                                                                                   1n
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               tomi;
                                                                                        MEDLINE=98196666; PubMed=9537320;
Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L., Graham D.E., Overbeek R., Snead M.A., Keller M., Aujay M., Huber Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
"The complete genome of the hyperthermophilic bacterium Aquifex
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                         +
                                                                                                                                                                                                                                                                                                                                                                                                                                   ß
                                                                                                                                                                                                      Nature 392:353-358(1998).

Nature 392:353-358(1998).

-!- CATALYTIC ACTIVITY: ATP + L-threonine + trnA(Thr) = AMP diphosphate + L-threonyl-trnA(Thr).

-!- COFACTOR: BINDS 1 ZINC ION (BY SIMILARITY).

-!- SUBUNIT: HOMODIMER (BY SIMILARITY).

-!- SUBCELLULAR LOCATION: Cytoplasmic.

-!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETAS
                                                                                                                                                                                                                                                                                                                                                             This SWISS-PROT entry is copyright. It is produced through a between the Swiss Institute of Bioinformatics and the EMB the European Bioinformatics Institute. There are no restriuse by non-profit institutions as long as its content is modified and this statement is not removed. Usage by and entities requires a license agreement (See http://www.isb-si) or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Euteleos
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CIZ1_HUMAN STANDARD; PRT; 898 AA.
Q9ULV3; Q9Y3G0; Q9Y3F9; Q9UHK4; Q9NYM8;
16-OCT-2001 (Rel. 40, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Cip1-interacting zinc finger protein (Nuclear protein NP94)
CIZ1 OR NP94 OR LSFR1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              638;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 9
  Bacteria; Aquificales; Aquificaceae; Aquifex.
NCBI_TaxID=63363;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Η.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score 51; DB Pred. No. 29; 6; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (ISOFORM 1).
PubMed=10529385
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1 MEKLFKEVKLEELENQKGSGLGK 23
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          26.4%;
47.8%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (Human)
                                                           FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A. MEDLINE-20001941;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Homo sapiens
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              11;
                                                                                                                                                                                             aeolicus.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CIZ1_HUMAN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Best
SFFFF WERE DRAWER DRAWE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ð
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                     conserved synteny chromosomal
                                                                                                                                                                                                                                                            AND 2/NP94B; ARE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                 CIP/WAF1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length 898;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MP94E

A NP94B)

2).

2).

2).

REF. 2).

REF. 2; AAF37882).

TREF. 2; AAF23237

TREF. 2; AAF237

TREF. 2; AAF37

TREF. 2; AAF37
                                                                                                                                                                                                        SUBCELLULAR LOCALIZATION OF
                                                                       Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases [3] SEQUENCE OF 649-898 FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      (IN ISOFORM NP94B)
(IN ISOFORM NP94B)
(IN REF. 2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Indel
                                                                                                                                                     of c
for
                                                                                                                                                                                                                        SUBUNIT: INTERACTS WITH CIP/WAF1.
SUBCELLULAR LOCATION: Nuclear.
ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE)
PRODUCED BY ALTERNATIVE SPLICING.
SIMILARITY: CONTAINS 1 MATRIN-TYPE ZINC FINGER.
                                               a human
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             C2H2_1; UNKNOWN_2.
Alternative splicing
                                                                                                                                             between the Fugu and human genomes: implications evolution and the cloning of disease genes."; Hum. Mol. Genet. 8:1313-1320(1999).

-!- FUNCTION: MAY REGULATE THE SUBCELLULAR INCALT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       28C7D76C6D4CC66B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB 1;
5
                             Warder D.E., Keherly M.J.; "NP94, a novel nuclear protein identified in
 AND
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (IN REF. (IN REF. (IN REF. (IN REF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 46;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        GLN-RICH.
GLN-RICH.
GLN-RICH.
GLU-RICH.
MISSING (IN I
Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 50.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                26
                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL; AB030835; BAA85783.1; -.
EMBL; AF159025; AAF23231.1; ALT_INIT.
EMBL; AF234161; AAF37882.1; ALT_INIT.
EMBL; Y17453; CAB44346.1; -.
EMBL; Y17454; CAB44347.1; -.
InterPro; IPR003604; ZnF_U1.
InterPro; IPR000690; ZnF_Matrin.
InterPro; IPR000822; ZnF_C2H2.
SMART; SM00355; ZnF_C2H2; 2.
SMART; SM00451; ZnF_U1; 3.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; U
FROM N.A. (ISOFORMS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                --VKLEELENQKGSGLGKAQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               X G C X D D C X D D X
                                                                                                       SEQUENCE OF 649-898 FROM N.A. MEDLINE=99299247; PubMed=10369878;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ^ ^ ^ <u>^</u>
                                                                                                                                     Gilley J., Fried M.; "Extensive gene order differences
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SONNSOORI
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               protein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            35
232
500
555
568
634
677
698
735
810
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   26.2%;
llarity 35.5%;
Conservative
            TISSUE=Medulloblastoma;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           44
470
830
761
120
201
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Nuclear
2-898
```

in

rsp.

```
EMBL; M35996;
HSSP; P03437;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       HEMA_IADA1
                                                                                                                       [1]
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                         CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                  CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P03443;
                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMA_IADA1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      7
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  uery
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     RESULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DT
DE
DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        3
                                                                                                                                                                                                                                                                  collaboration
                                                                                                                                                                                                                                                                                                         commercial
                                                                                                                                                                                                                                                                                                                sib.ch/announce,
                                                                                                                                                                                                                                                                            outstation
                                                                                                                                                                                                                                                                            ne EMBL outstat:
restrictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                              gene families in
                                                                                                                                                                                                                                                                                                                                                                                                                                                   binding;
                                                                                                    Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                              ZINC
                                                                                                                                                                                                                                                                                                         for
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         ••
                                                                                                                                                                                                                                                                                                                                                 HSSP; P15822; 1BBO.
InterPro; IPR001909; KRAB.
InterPro; IPR000822; Znf-C2H2.
Pfam; PF01352; KRAB; 1.
Pfam; PF00096; Zf-C2H2; 8.
SMART; SM00349; KRAB; 1.
SMART; SM00355; ZnF_C2H2; 8.
PROSITE; PS50805; KRAB; 1.
PROSITE; PS50805; ZINC_FINGER_C2H2_1; 8.
PROSITE; PS50157; ZINC_FINGER_C2H2_2; 10.
Transcription regulation; DNA-binding; Zinc-finger; Metal-binc
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     1;
                                                                                                                                                                                                                                                                    ø
                                                                                                                                                                            "Differential expansion of homologous zinc-finger gene famil human chromosome 19q13.2 and mouse chromosome 7.";
Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
-!- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE FINGER PROTEINS.
-!- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
                                                                                                                                                                                                                                                                                                       and
                                                                                                                                                                                                                                                               This SWISS-PROT entry is copyright. It is produced through between the Swiss Institute of Bioinformatics and the Enthe European Bioinformatics Institute. There are no restruse by non-profit institutions as long as its content modified and this statement is not removed. Usage by ancentities requires a license agreement (See http://www.isb-sor send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      45
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indels
                                                                                                                                                             Gordon L., Ashworth
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           -CAALWLQCAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        13;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   55F55A778116D184
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ij
                                              (Rel. 40, Created)
(Rel. 40, Last sequence update)
(Rel. 40, Last annotation update)
protein 222.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Score 49.5; D
Pred. No. 33;
4; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ZINC FINGERS.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
C2H2-TYPE.
                                                                                                                                                             M., Branscomb E., Hauser L.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                        4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           KLFKEVKLEELENQKGSGLGKAQ-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    25.6%;
ilarity 30.6%;
Conservative
                                                                                                                                                                                                                                                                                                                                           EMBL; AF187988; AAF04104.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            STANDARD;
                                     STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
15; Conser
                                                                                                                                                    FROM N.A
                                                                                                                                  NCBI_TaxID=9606;
                                                O9UK12;
16-OCT-2001
16-OCT-2001
16-OCT-2001
                                                                                Zinc finger
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            HEMA_IAMAA
                                     2222_HUMAN
Q9UK12;
                                                                                                                                                           Shannon M. Stubbs L.;
                                                                                                                                                    SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Local
                                                                                           ZNF222
                          1222_HUMAN
ID 2222_H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HEMA_IAMAA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           က
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         RESULT
                   RESULT
                                                                                                                                                           Π
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ò
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TWO CHAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          VIRUS
                                                                                                                                                                                                                         Ø
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                       (Influenza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           ж
;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      influenza A virus.";

Virology 179:759-767(1990).

-! FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE V
CELL RECEPTORS AND FOR INITIATING INFECTION.

-! SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWC
(HA1 AND HA2) LINKED BY A DISULFIDE BOND.

-! SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         563
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Webster
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      subtype
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                chain;
                                                                                                                         chain
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Influenza A virus (strain A/Duck/Alberta/28/76).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
                                                                                                                                                                           virus (strain A/Mallard/Gurjev/244/82)).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae; Influenza virus A and B group; Influenza A viruses.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CHAIN.
CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           21-JUL-1986 (Rel. 01, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HAl
P26137;
01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HA1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 MEDLINE=91049442; PubMed=2238469;
Kawaoka Y., Yamnikova S., Chambers T.M., Lvov D.K.,
"Molecular characterization of a new hemagglutinin,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HEMAGGLUTININ HA1 CHEMAGGLUTININ HA2 CN-LINKED (GLCNAC.N-LINKED (GLCNAC.N-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      FBD2BA976A3D5977
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL; masser; HTM.
HSSP; P03437; 1HTM.
InterPro; IPR001364; Hemagglutn.
Pfam; PF00509; Hemagglutinin; 1.
ProDom; PD000225; Hemagglutn; 1.
Envelope protein; Hemagglutinin; Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         564 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     486
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ed. No. 40;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score 49.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                MEKLFKEVKLEELENOKGSGLGKAQCAALWLQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   ---CFEIFHQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -; NOT_ANNOTATED_CDS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              62590 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         25.6%;
36.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        457 MNKLFERVRROLRENAEDOGNG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            341
16
16
33
177
177
237
309
496
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           <1
13
343
16
33
177
237
237
309
496
563 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Similarity
12; Conser
                                                                                                                                                                                                                                                                                                                                                                                               NCBI_TaxID=11433;
```

14:40:04

4

2

Oct

rsp.

37

Н

coby

-030 - 3

```
collaboration
                                                                                                                                                                                                                                                                                                                                   n no way
commercial
                                                                                                                                                                                                                                                                                      This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          ij
                                                                                                                                                                                                                                            CHAINS
                                                                                                                                                                               subtypes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Gaps
                                                                                                                                                                                                                   VIRUS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          (POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL). (POTENTIAL).
                                                                                                                                                                                                                                            8
                                                                                                                                                                                                 Proc. Natl. Acad. Sci. U.S.A. 78:7639-7643(1981).

-!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VELL RECEPTORS AND FOR INITIATING INFECTION.

-!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWC (HAI AND HA2) LINKED BY A DISULFIDE BOND.

-!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
                                            SEQUENCE FROM N.A.
MEDLINE=89204912; PubMed=2705304;
Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.;
"Distinct lineages of influenza virus H4 hemagglutinin genes different regions of the world.";
Virology 169:408-417(1989).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Length 564;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         A/Ruddy Turnstone/New Jersey/47/85 trand viruses; Orthomyxoviridae;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Indel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        CRC64;
                                                                                                                                                                             genes
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CHAIN.
CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HAl
   viruses;
                                                                                                                                                                           Sequence relationships among the hemagglutinin f influenza A virus."; roc. Natl. Acad. Sci. U.S.A. 78:7639-7643(1981)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     808C4BE89E15C94D
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DB 1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   HEMAGGLUTININ HA1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           N-LINKED (GLCNAC N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                N-LINKED (GLCNAC
N-LINKED (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (GLCNAC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        (IN REF
  group; Influenza A
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         487
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A virus (strain A/Ruddy Turnst ssRNA negative-strand viruses;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Score 49.5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             MEKLFKEVKLEELENQKGSGLGKAQCAALWLQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Pred.
                                                                                                                                     SEQUENCE OF 1-101 FROM N.A. MEDLINE=82150925; Pubmed=6174976;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Hemagglutn; 1.
Hemagglutinin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                PRINTS; PR00509; Hemagglutinin; 1. ProDom; PD000225; Hemagglutinin; 1. Envelope prototics.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       69 F
63381 MW;
                                                                                                                                                                                                                                                                                                                                                                                         EMBL; M25287; AAA43220.1; -. EMBL; J02102; AAA43179.1; -. PIR; E34214; HMIVF5.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             25.6%;
36.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Conservative
  and B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16
342
564
18
34
178
310
                                                                                                                                                                                                                                                                                                                                                                                                                                           P03437; 2HMG.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      ¥;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity 12; Conser
Influenza virus A
Influenza A virus,
NCBI_TaxID=11346;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             344
18
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      34
178
310
497
69
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Influenza A
Viruses; ssl
                                                                                                                                                                Air G.M.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CARBOHYD
CONFLICT
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               P19700;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       IARUD
                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              uery
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             SULT
                                               a
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AC
DT
DT
DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             OC OC
```

```
1;
                                                                                                                                                           CHAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          CHAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=89204912; PubMed=2705304;
MEDLINE=89204912; PubMed=2705304;
Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.;
Toistinct lineages of influenza virus H4 hemagglutinin genes in different regions of the world.";
Virology 169:408-417(1989).

-!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS CELL RECEPTORS AND FOR INITIATING INFECTION.

-!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAI (HAI AND HA2) LINKED BY A DISULFIDE BOND.

-!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
                                                                                                                                 VIRUS
                                                                                                                                                                                                                                                                                                                            (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                         TWO
                                                                                                                                                                                                                                                                                                                                                                                                                                                 3,
                                                                                                  FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE CELL RECEPTORS AND FOR INITIATING INFECTION.
SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO (HAI AND HA2) LINKED BY A DISULFIDE BOND.
SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY (BA); P03437; 24MG.
                                                                                          genes
                                                                                                                                                                                                                                                                                                                                                                                                                       564
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Influenza A virus (strain A/Seal/Massachusetts/133/82).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza virus A and B group; Influenza A viruses;
Influenza A virus.
NCBI_TaxID=11460;
                                                                                                                                                                                                                                                                                                                                                                                                                                                 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                       Length
                                    Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.; "Distinct lineages of influenza virus H4 hemagglutinin different regions of the world."; virology 169:408-417(1989).
                                                                                                                                                                                                                                                                                                                                                                                             CRC64;
                                                                                                                                                                                                                                                                             Signal
                                                                                                                                                                                                                                                                                                      CHAIN.
CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HA1
Hemagglutinin HA2 chain].
  viruses;
                                                                                                                                                                                                                                                                                                  HEMAGGLUTININ HA1 CE
HEMAGGLUTININ HA2 CH
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                12
                                                                                                                                                                                                                                                                                                                                                                                                                       1;
                                                                                                                                                                                                                                                                           Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                       DB
   K
                                                                                                                                                                                                                                                                                                                                                                                                                                     40;
                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    487
  Influenza
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  564
                                                                                                                                                                                                                                                                                                                                                                                                                       Score 49.5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEKLFKEVKLEELENOKGSGLGKAQCAALWLQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        InterPro; IPR001364; Hemagglutn.
Pfam; PF00509; Hemagglutinin; 1.
PRINTS; PR00329; HEMAGGLUTN12.
ProDom; PD000225; Hemagglutn; 1.
Envelope protein; Hemagglutinin; G.
SIGNAL 17 342 HEMAGCCHAIN 344 564 HEMAGCCARBOHYD 18 18 18 N-LINICARBOHYD 178 178 N-LINICARBOHYD 310 310 N-LINICARBOHYD 310 N-LINICARBOHYD 310 N-LINICARBOHYD 310 N-LINICARBOHYD 310 N-LINICARBOHYD 178 178 N-LINICARBOHYD 310 N-LINICARBOHYD 1497 497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Pred
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            P03437; 2HMG.
Pro; IPR001364; Hemagglutn.
PF00509; Hemagglutinin; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                9
   group;
                                                                                                                                                                                                                                                                                                                                                                                            63271 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                       25.6%;
36.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  STANDARD;
  B
  and
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              I34214; HMIVF9.
P03437; ZHMG.
                                                                                                                                                                                                                                                                                                                                                                                             AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Similarity 12; Conser
           Influenza A virus
NCBI_TaxID=11458;
  virus
                                                                                                                                                                                                                                                                                                                                                                                             64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               HEMA_IASE2
P19701;
01-FEB-1991
01-FEB-1991
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                       Query Match
                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Pfam; PF
PRINTS;
                                                                                                                                                                                                                                                                                                                                                                                                                                    Local
                                                                                                                                                                                              PIR; (
HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    HEMA_IASE2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PIR; ]
HSSP;
                                                                                                                                                                                                                                                                                                                                                                                                                                     Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 458
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SULT
g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  δ
```

CRC64

A8A3D998BBE3A58F

63249 MW;

¥,

SEQUENCE

```
qq
                                                                                                                                                                                                                                               DE
DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          collaboration -
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   n no way
commercial
                                                                                                                                                                                             ij
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | for commercial
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          gh a collabora
EMBL outstati
strictions on
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       TWO CHAINS
                                                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            VIRUS
                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.

SEQUENCE FROM N.A.

MEDLINE-89204912; PubMed-2705304;

MEDLINE-89204912; PubMed-2705304;

Donis R.O., Bean W.J., Kawaoka Y., Webster R.G.;

I "Distinct lineages of influenza virus H4 hemagglutinin genes in different regions of the world.";

Virology 169:408-417(1989).

CELL RECEPTORS AND FOR INITIATING INFECTION.

CELL RECEPTORS AND FOR INITIATING INFECTION.

CELL RECEPTORS AND FOR INITIATING INFECTION.

CHAI AND HA2) LINKED BY A DISULFIDE BOND.

CHAIL AND HA2

CHAIL AND H
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    TAL).
TAL).
TAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                LY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ( POTENTI ( ) ( POTENTI ( ) ( POTENTI ( ) ( POTENTI ( ) ) ( POTENTI ( ) )
                                                                    . .) (POTENT]
                                                                                                                                                                 564
                                                                                                                                                                                                                                                                                                                                                                                       chain;
                                                                                                                                                                                                                                                                                                                                                                                       Lustuenza A virus (strain A/Turkey/Minnesota/833/80).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
Influenza virus A and B group; Influenza A viruses;
Influenza A virus.
NCBI_TaxID=11468;
                                                                                                                                                                                            Indels
                                                                                                                                                                 Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Signal
                Signal
                                         CHAIN.
                                                                                                                                                                                                                                                                                                                                         01-FEB-1991 (Rel. 17, Created)
01-FEB-1991 (Rel. 17, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HA1
Hemagglutinin HA2 chain].
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             CHAIN.
                                                                                                                                                                                          12;
                                                                 N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
N-LINKED (GLCNAC...
                                                                                                                       N-LINKED (GLCNAC. .. 455D17B1ADD2B82B
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          HEMAGGLUTININ HA1 (HEMAGGLUTININ HA2 (N-LINKED (GLCNAC.N-LINKED (GLCNAC.N-LINKED (GLCNAC.N-LINKED (GLCNAC.N-LINKED (GLCNAC.N-LINKED (GLCNAC.
                                                                                                                                                                DB 1;
               Glycoprotein;
                                         HEMAGGLUTININ HA1
HEMAGGLUTININ HA2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Glycoprotein;
                                                                                                                                                                           40;
                                                                                                                                                                                          Mismatches
                                                                                                                                                                                                                                                487
                                                                                                                                                              49.5;
                                                                                                                                                                                                                                    458 MNKLFERVRRQLRENAEDKGNG---CFEIFHQC
                                                                                                                                                                                                                    MEKLFKEVKLEELENQKGSGLGKAQCAALWLQC
                                                                                                                                                              Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Hemagglutn; 1.
Hemagglutinin; 1.
16
342 HEMA(
564 HEMA(
18 N-LII
34 N-LII
178 N-LII
310 N-LII
3497 N-LII
Hemagglutn; 1.
Hemagglutinin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    InterPro; IPR001364; Hemagglutn. Pfam; PF00509; Hemagglutinin; 1. PRINTS; PR00329; HEMAGGLUTN12. ProDom; PD000225; Hemagglutn; 1. Envelope protein; Hemagglutinin;
                                                                                                                                                                                          : 9
                        16
342
H
564
H
18
N
34
N
178
N
310
N
497
N
                                                                                                                                                                25.6%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EMBL; M25290; AAA43223.1;
PIR; H34214; HMIVF8.
                                                                                                                                                                                            Conservative
                                                                                                                                                                                                                                                                                                                  STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              2HMG.
                                                                                                                                       AA;
                                                                                                                                                                             Similarity
12; Conser
              Envelope protein; SIGNAL 1
                                    17
344
18
34
178
310
497
564
 PD000225
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        344
118
18
34
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 178
310
497
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             HSSP; P03437;
                                                                                                                                                                                                                                                                                                                 HEMA_IATKM
P19702;
                                                                 CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                 Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 CARBOHYD
                                                                                                                                                                           Local
 ProDom;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SIGNAL
                                                                                                                                                                                                                                                                                      KESULT 12
HEMA_IATKM
                                       CHAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CHAIN
                                                                                                                                                                                                                    ᠳ
                                                                                                                                                                                         Matches
                                                                                                                                                                              Best
                                                                                                                                                                                                                                                                                        RESULT
DR
KW
FT
FT
FT
FT
FT
                                                                                                                                                                                                                                                                                                                                                                                  q
                                                                                                                                                                                                                                                                                                                                          PP PP
                                                                                                                                   S
                                                                                                                                                                                                                     ò
```

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         1;
                                          ä
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   TWO CHAINS
                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                      SEQUENCE FROM N.A.

MEDLINE=91049442; PubMed=2238469;

Kawaoka Y., Yamnikova S., Chambers T.M., Lvov D.K., Webster R.G.;

Kawaoka Y., Yamnikova S., Chambers T.M., Lvov D.K., Webster R.G.;

Kawaoka Y., Yamnikova S., Chambers T.M., Lvov D.K., Webster R.G.;

Kawaoka Y., Yamnikova S., Chambers T.M., Lvov D.K., Webster R.G.;

Influenza A virus.";

Virology 179:759-767(1990).

-!- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS

CELL RECEPTORS AND FOR INITIATING INFECTION.

-!- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAI

(HAI AND HA2) LINKED BY A DISULFIDE BOND.

-!- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
                                                                                                                                                                                                                                                                                                                               Ø
                                                                                                                                                                                                                                                                                                    uniuenza A virus (strain A/Mallard/Astrakhan/263/82) (Influenza virus (strain A/Mallard/Gurjev/263/82)).
Viruses; ssRNA negative-strand viruses; Orthomyxoviridae; Influenza virus A and B group; Influenza A viruses;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                         ä;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        568;
                                                                                                                                                                                                                       01-MAY-1992 (Rel. 22, Created)
01-MAY-1992 (Rel. 22, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Hemagglutinin precursor [Contains: Hemagglutinin HAI chain;
Hemagglutinin HA2 chain].
                                           Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Length
            Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CRC64
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Signal
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CHAIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 HEMAGGLUTININ HA1 CH
HEMAGGLUTININ HA2 CF
N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        12;
           1;
                                         12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Glycoprotein;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB
            DB
               ..
40;
                                          Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Mismatches
                                                                                                            487
                                                                           33
                                                                                                                                                                                             568
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Score 49.5;
Pred. No. 4
            S.
                                                                           MEKLFKEVKLEELENQKGSGLGKAQCAALWLQC
                                                                                                           MNKLFERVRROLRENAEDKGNG---CFEIFHQC
           Score Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  M35997; -; NOT_ANNOTATED_CDS. A46339; A46339.
                                                                                                                                                                                             PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                PIR; A46339; A46339.
HSSP; P03437; 1HTM.
InterPro; IPR001364; Hemagglutnin; 1.
Pfam; PF00509; Hemagglutinin; 1.
PRINTS; PR00329; HEMAGGLUTN12.
ProDom; PD000225; Hemagglutn; 1.
Envelope protein; Hemagglutinin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Hemagglutn; 1.
Hemagglutinin;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           9
                                          9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                346
568
21
21
83
62
N
182
N
242
N
314
N
501
N
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       25.6%;
36.4%;
           25.6%;
36.4%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Conservative
                                           Conservative
                                                                                                                                                                                              STANDARD;
                       1 Similarity
12; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                NCBI_TaxID=11434;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              18
348
21
38
62
182
242
314
501
                                                                                                                                                                                             HEMA_IAMAB
P26136;
01-MAY-1992 (
01-MAY-1992 (
16-OCT-2001 (
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      CARBOHYD
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
Query Mac.
             Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SIGNAL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    EMBL;
                                                                                                                                                                           IAMAB
                                                                           -
                                                                                                            458
                                            Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Best
```

33

MEKLFKEVKLEELENQKGSGLGKAQCAALWLQC

,-

CRC64;

369593A1F0B558C8

84998 MW;

754 AA;

SEQUENCE

```
collaboration
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                                                                                                                           (Aspartate beta-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESIDUE IN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 kDa (AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       + 0(2)
                                                                                                                                                                                                                                                                  SEQUENCE FROM N.A.

TISSUE=Liver, and Brain;

MEDLINE=92332546; PubMed=1378441;

Jia S., Vandusen W.J., Diehl R.E., Kohl N.E., Dixon R.A.F.,

Elliston K.O., Stern A.M., Friedman P.A.;

"cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-
hydroxylase.":
                                                                                                                                                                                  Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoldea;
Bovidae; Bovinae; Bos.
NCBI_TaxID=9913;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Iron; Transmembrane; Signal-anchor;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SUBUNIT: MONOMER.
SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    2-oxoglutarate + CO(2).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 PTM: MIGHT BE PROCESSED TO THE 56 kDa (AA 289-754) OR 52 311-754) FORMS IN THE LUMEN OF THE ENDOPLASMIC RETICULUM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              (POTENTIAL).
R (TYPE-II MEMBRANE
                                                                                                  01-NOV-1997 (Rel. 35, Created)
01-NOV-1997 (Rel. 35, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
Aspartyl/asparaginyl beta-hydroxylase (EC 1.14.11.16) (Aspanydroxylase) (ASP beta-hydroxylase) (Peptide-aspartate beta-
                                                                                                                                                                                                                                                                                                                                                                                                                                  ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                  Garsky V.M., Stern
                                                                                                                                                                                                                                                                                                                                                                                                                                                        "Bovine liver aspartyl beta-hydroxylase. Purification and characterization.";
J. Biol. Chem. 266:14004-14010(1991).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              IN ASP OR ASN F
(EGF) DOMAINS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      222
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           (FOIENTIAL).
LUMENAL (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    (GLCNAC.
(GLCNAC.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   CATALYTIC ACTIVITY: Peptide L-aspartate + peptide 3-hydroxy-L-aspartate + succinate COFACTOR: IRON.
                                                                           754 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SIGNAL-ANCHOR
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             FUNCTION: SPECIFICALLY HYDROXYLATES AN CERTAIN EPIDERMAL GROWTH FACTOR-LIKE (
            491
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CYTOPLASMIC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     (POTENTIAL)
                                                                                                                                                                                                                                                                                                                                                                                                                      ', C.J.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         POLY-SER.
POLY-LYS.
N-LINKED
462 MNKLFERVRRQLRENAEDQGNG---CFEIFHQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 N-LINKED
N-LINKED
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             POLY-GLY
                                                                                                                                                                                                                                                                                                                                                                 Chem. 267:14322-14327(1992)
                                                                                                                                                                                                                                                                                                                                                                                                                  PubMed=1856229;
W.J., Petroski
                                                                                                                                                                                                                                                                                                                                                                                          OF 289-385 AND 615-641.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Oxidoreductase; Dioxygenase;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AAA03563.1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          TPR
                                                                           STANDARD;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            57
78
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 754
12
21
328
96
466
702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Endoplasmic reticulum
DOMAIN 1 5
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        PF00515; TPR; 2
                                                                                                                                                                                                                                                                                                                                                                                                                MEDLINE=91310689;
Wang Q., Vandusen
Friedman P.A.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 7
28
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   99
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        14
318
96
466
702
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      reticulum.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              M91213;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         PROTEINS.
                                                                                                                                                                                                                                                                                                                                                                                                       TISSUE=Liver;
                                                                       ASPH_BOVIN
Q28056;
01-NOV-1997
01-NOV-1997
                                                                                                                                                                                                                                                                                                                                                   hydroxylase.
J. Biol. Che
                                                                                                                                                                     dioxygenase)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          InterPro;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DOMAIN
DOMAIN
CARBOHYD
CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        TRANSMEM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               DOMAIN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              EMBL;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Pfam;
                                                 RESULT
                                                                                                                qq
```

its

```
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
                                  1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    MEDLINE=91358413; PubMed=1885571;
Pei Z., Ellison R.T. III, Blaser M.J.;
"Identification, purification, and characterization of major
antigenic proteins of Campylobacter jejuni.";
J. Biol. Chem. 266:16363-16369(1991).
-!- FUNCTION: COMMON ANTIGEN AND A MAJOR CELL ADHERENCE MOLECULE. MOST
PROBABLY INVOLVED, WITH PEBIC, IN A BINDING-PROTEIN-DEPENDENT
TRANSPORT SYSTEM FOR AN AMINO ACID. MAY BE INVOLVED IN BINDING TO
INTESTINAL CELLS.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SUBCELLULAR LOCATION: CELL SURFACE.
SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    tor of Campylobacter jejuni, is a in Gram-negative nutrient transport
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             jejuni
                                                                                                                                                                                                                                                                                                       Campylobacter jejuni.
Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
Campylobacter.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 STRAIN=NCTC 11168;
MEDLINE=20150912; PubMed=10688204;
Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C., Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W., Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M., Whitehead S., Barrell B.G.;
"The genome sequence of the food-borne pathogen Campylobacter j reveals hypervariable sequences.";
Nature 403:665-668(2000).
                                  11;
Length 754;
                                  Indels
                                                                                                                                                                                                                       01-NOV-1995 (Rel. 32, Created)
01-NOV-1995 (Rel. 32, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
Major cell-binding factor precursor (CBF1) (PEB1).
                                  9
   DB 1;
                                                                                                      380
                                                                    26
                   ed. No. 53;
Mismatches
 Score 49.5;
Pred. No. 53;
                                                                                      -QKGSGLGKAQC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                   of
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 or send ...

EMBL; L13662; AAA02919.1; -.

EMBL; AL139076; CAB73178.1; -.

InterPro; IPR001638; SBP_bac_3.

TherPro; IPR001311; SBP_glut_receptor.

TherPro; IPR001311; SBP_glut_receptor.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     268:18717-18725(1993).
                                                                                                                                                                                                                                                                                                                                                                                                                                          Pei Z., Blaser M.J.; "PEB1, the major cell-binding factor homolog of the binding component in G systems.";
                                                                                                                                                                                                                                                                                                                                                                                             SEQUENCE FROM N.A.
STRAIN-ATCC 55026 / 81-176;
MEDLINE-93366784; PubMed=8360165;
Similarity 38.98; 14; Conservative
                                                                                                                                                                                            STANDARD;
                                                                    EKLFKEVKLEELEN-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            27-53.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PROTEIN FAMILY
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                                                                 NCBI_TaxID=197;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Chem.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 OF.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE
   Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Biol.
                    Local
                                                                                                                                                                     CAMJE
                                                                                                      345
                                                                    ~
                                                                                                                                                          RESULT 15
                      Best
                                                                                                                                                                                                                                                                           GN
                                                                                                      g
                                                                                                                                                                                                                                                                                                                                                                               RP
RX
RA
RA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          RA
RA
RT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RI
RN
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ŏ
```

```
DR SMART; SM00062; PBPb; 1.

DR . PROSITE; PS01039; SBP_BACTERIAL_3; 1.

KW Signal; Transport; Antigen; Complete proteome.

FT SIGNAL 1 26

FT CHAIN 27 259 MAJOR CELL-BINDING FACTOR.

SQ SEQUENCE 259 AA; 28177 MW; DA132ECE87BBCBD5 CRC64;

Query Match

Query Match

A1.7%; Pred. No. 23;

Best Local Similarity 41.7%; Pred. No. 23;

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 4 LFKEVKLEELENQKGSGLGKAQCA 27

Cy 4 LFKEVKLEELENQKGSGLGKAQCA 27

Db 134 VLKEKKYKSLADMKGANIGVAQAA 157
```

Search completed: October 24, 2002, 09:28:23 Job time : 7.32117 secs

·	
	•
	•
	, • • • • • • • • • • • • • • • • • • •

J

```
updates/sec
                                                              Seconds
                                                             ; Search time 11.8832 (without alignments) 538.643 Million cell
                                                                                                                            37
5.1.3
Compugen Ltd
                                                                                                      US-09-894-030-3_COPY_1_37
193
1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                                                                                                               562222 seqs, 172994929 residues
version
- 2002
                                          model
                                                              2002, 09:24:21
                                                                                                                                                           Ŋ
                                                                                                                                                            0
                                          30
GenCore
(c) 1993
                                                                                                                                                           Gapext
                                          using
                                          protein search,
                                                             24,
                                                                                                                                                 BLOSUM62
Gapop 10.0
           Copyright
                                                              October
                                                                                                                                                  table:
                                                                                                                   score:
                                          OM protein
                                                                                                                              Sednence:
                                                                                                                                                                                 Searched:
                                                                                                                  Perfect
                                                              Run on:
                                                                                                                                                  Scoring
                                                                                                        Title:
```

562222

S

satisfying chosen parameter

hits

of

number

Total

0 2000000000

seq length: seq length:

DB DB

Minimum

Maximum

Post-processing: Minimum Match 08

Maximum Match 100% Listing first 45 summaries SPTREMBL\_19 Listing Database

sp\_archea:\*
sp\_bacteria:\*
sp\_fungi:\*
sp\_human:\*
sp\_invertebrate:\*
sp\_mammal:\*
sp\_mammal:\*
sp\_organelle:\*
sp\_phage:\*
sp\_phage:\*
sp\_rodent:\*
sp\_rodent:\*
sp\_virus:\*
sp\_virus:\*
sp\_virus:\*
sp\_virus:\*
sp\_vartebrate:\*
sp\_vartebrates:\*
s 

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

064033 bacteriopha
034781 bacillus su
002290 neocallimas
094f34 arabidopsis
095102 arabidopsis
0902w3 caenorhabdi
099n24 drosophila
099n13 drosophila
099n05 drosophila
099n303 homo sapien
09h304 homo sapien
09h304 homo sapien
09h099 homo sapien
09ng74 homo sapien
09ng22 caprine art
095u08 drosophila Description 064033 034781 Q02290 Q94F34 Q95J02 Q9GN24 Q9GN13 Q9GN05 Q9H304 Q9H304 Q9H304 Q9HQQ9 Q9HQQ9 Q9HQQ9 DB Length 100.0 100.0 31.6 29.0 28.2 28.2 28.2 28.2 28.2 28.2 28.2 Query Match Score No. Result

080181 streptococc 034073 streptococc Q92487 aeromonas p Q992p6 mus musculu Q91wj2 oryza sativ Q91wj5 oryza sativ Q9asw9 arabidopsis 082634 arabidopsis	3um9 helicobacte 3um8 helicobacte 3um6 helicobacte 97zb8 sulfolobus 93xb2 rosa hybri gtul drosophila	Q91wg8 oryza sativ Q9fxil arabidopsis Q91920 caenorhabdi Q21920 caenorhabdi Q911c0 pseudomonas Q914d6 xanthomonas Q937u6 borrelia he Q937u6 borrelia he	una3 schizosacch 9si41 arabidopsi 9795 schizosacch btg3 homo sapien h868 homo sapien 59083 pyrococcus
53 27.5 1000 9 080 53 27.5 1006 9 034 2.5 27.2 729 2 092 2.5 27.2 1322 11 09 52 26.9 475 10 09 52 26.9 713 10 09	52 26.9 1287 2 Q93UM9 52 26.9 1288 2 Q93UM8 52 26.9 1289 2 Q93UM6 1.5 26.7 295 17 Q97ZB 1.5 26.7 443 10 Q93XB 1.5 26.7 453 5 Q9GTU1	1.5 26.7 811 10 Q9FWG 1.5 26.7 1024 10 Q9FXI 1.5 26.7 2584 5 Q9TW88 1.5 26.7 2606 5 Q21920 51 26.4 361 16 Q9IIC 51 26.4 368 2 Q9L4D6 51 26.4 514 2 Q937U6 51 26.4 526 10 Q9LMD	1 26.4 558 3 0 1 26.4 1738 10 26.2 633 3 0 5 26.2 633 3 0 5 26.2 837 4 0 5 26.2 842 4 0 0 25.9 148 17
17 18 19 20 22 23 23	.00840°		v0112647

## ALIGNMENTS

Gaps SPbetac2 Siphoviridae ö : 56; "The complete nucleotide sequence of the Bacillus subtilis prophage.";
Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
EMBL; AF020713; AAC12992.1; -.
Lipoprotein.
SEQUENCE 56 AA; 5982 MW; 79EC0BF822F9F4C0 CRC64; SEQUENCE FROM N.A. Lazarevic V., Duesterhoeft A., Soldo B., Hilbert H., Mauel Karamata D.; Indels Length Caudovirales; Last sequence update) Last annotation update) Score 193; DB 9; Pred. No. 2.4e-19; ; Mismatches 0; 1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG 37 56 AA stage; Created) PRT; RNA 100.0%; s. larity 100.0%; F. Conservative 0; 20 07, 07, 19, 064033; 064033; 01-AUG-1998 (TrEMBLrel. 0 01-AUG-1998 (TrEMBLrel. 0 01-DEC-2001 (TrEMBLrel. 1 PUTATIVE LIPOPROTEIN. PRELIMINARY; viruses, Bacteriophage SPBc2 Viruses; dsDNA viru NCBI\_TaxID=66797; Similarity 37; Conserv Query Match Best Local Best Loc Matches YOLG RESULT 064033 G δ

1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG d

ö

Last sequence update)
Last annotation update) 56 AA. Created) PRT; 05, 05, 08, (TrEMBLrel. (TrEMBLrel. (TrEMBLrel. PRELIMINARY; O34781; O34781; O1-JAN-1998 ( O1-JAN-1998 ( O1-NOV-1998 ( 7 RESULT 034781 GI

.rspt

37

```
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            REPEAT
REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           EPEAT
EPEAT
EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       EPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   EPEAT
    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                               in A.;
Bacillus
                                                                                                                                                          Must F., Oggaswars N., Moszer I., Albertini A.M., Alloni G.,
Azevedo V., Bertero N.G., Bessieres P., Bolotin A., Borchert S.,
Azevedo V., Bertero N.G., Bessieres P., Bolotin A., Borchert S.,
Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.M.,
Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
Entiz C., Fujita M., Fabret C., Ferrari E., Foulger D.,
Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
Chiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
Hilbert H., Holsappel S., Hosono S., Hullo W.F., Itaya M., Jones L.,
Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
A Joris B., Karamata A., Lardinois S., Mauel C., Mediue C.,
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
Tosato V., Uchlyama S., Vandenbol M., Vannier F., Vassarotti A.,
Tosato Y., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
"The complete genome sequence of the gran-positive bacterium Bacillus
                                                                                                                                                                                                                                                                     nemant A.,
., Jones L.,
Klein C.,
                                                                                                                                                                                                                                                                                                                                                                                               B
                                                                                                                                                                                                                                                                                                                             Noback M.,
Park S.H.,
tt A.M.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Gaps
                                                                                                                                                                          t S.,
Bron S.,
N.M.,
el R.A.,
son P.T.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                    ata K.,
                                                                                                                                                                                                                                                                                                                                                                                               Soldo
                                                                                                                                                                                                                                                                                                                                                        A.M.,
ynolds 9
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 56;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        A.;
databases.
                                                                                                Paik S.H., Hansen J.N.;
Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           79EC0BF822F9F4C0 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        002290;
01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-NOV-2001 (TrEMBLrel. 17, Last annotation update)
01-VIN-2001 (TrEMBLrel. 17, Last annotation update)
ENDO-1,4-BETA-XYLANASE B (EC 3.2.1.8) (XYLB).
                                Bacteria; Firmicutes; Bacillus/Clostridium group;
Bacillus/Staphylococcus group; Bacillus.
NCBI_TaxID=1423;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                193; DB 16;
No. 2.4e-19;
matches 0;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Kunst F., Ogasawara N., Yoshikawa H., Danchin
Submitted (NOV-1997) to the EMBL/GenBank/DDBJ
EMBL; AF014938; AAC63531.1; -.
EMBL; Z99115; CAB14066.1; -.
Complete proteome.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          37
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    860 AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          1 MEKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRT;
                                                                                                                                         STRAIN=168;
MEDLINE=98044033; PubMed=9384377;
 SUBLANCIN 168 PRECURSOR PEPTIDE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ;
0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               100.0%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          5982 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                subtilis.";
Nature 390:249-256(1997).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                         SEQUENCE FROM N.A.
STRAIN=168;
Paik S.H., Hansen
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Local Similarity
les 37; Conserv
                                                                                                                                FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  FROM N.A.
                     Bacillus subtilis
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            56 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                SEQUENCE FR
STRAIN=168;
                                                                                                                               SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    002290
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       002290
ð
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             AC
DT
DT
DT
DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    П
```

```
OG NGCOLILIMENTIN PATTICIATUM (Sumen fungus).

OC NGCOLILIMENTIN PATTICIATUM (Sumen fungus).

REPROSIDENCE FROM N.A.

REPROSID
```

3

.rspt

```
13;
                                                                                                                                                                                                                                                           Query Match
Best Local (
                                                                                                                                                                                                                                                                                     4
                                                                                                                                                                                                                                                                       Matches
                                                                                                                                                                                                                                                                                                                    9
                                                                                                                                                                                                                                                                                                                    SULT
        RESULT
Q9SJ02
                                                                                                                                                                                                                                                                                                                           Q9N2W3
                                              던
                                                                                                                                                                                                                                                                                                 g
                                                                                                                                                                                                                                                                                    ŏ
                                                                                                                                                                                                                                                                                                                                                                                                                                              1;
                                                                                                                                                                                                               ö
                                                                                                                                                                                                                                                                                                                    Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
eurosids II; Brassicales; Brassicaceae; Arabidopsis.
                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                                                                                                               Gaps
                                                                                                                                                                                                                                                                                                                                                                        H.,
Kawai
                                                                                                                                                                                                                                                                                                                                                                                       aki K.,
                                                                                                                                       (POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
(POTENTIAL).
                                                                                                                                                                                                                                                                                                                                                                 а М.,
                                                                                                                                                                                                                                                                                                                                                                                                                                               .;
;
                                                                                                                                                                                                              ô
                                                                                                                                                                                                                                                                                                                                                         SEQUENCE FROM N.A.

Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H. Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinoza Ecker J., Theologis A., Davis R.W.;

Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AF386927; AAK62372.1; -.

SEQUENCE 170 AA; 18220 MW; BD0FFB28CD3C01ED CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                  Length 170
                                                                                                                           (POTENTIAL)
                                                                                                                                                                                                  Length 860;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                                                                                                                               Indels
                                                                                                                                                                                     CRC64;
                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                      N-LINKED (GLCNAC...
                                                                                                                                                                                                                                                                                                                                                                                                                                              10;
                                                                                                                                                                                                              20;
39.
40.
41.
42.
43.
44.
45.
46.
49.
50.
51.
52.
53.
53.
55.
55.
55.
                                                                                                                                                                                                                                                                                                                                                                                                                                 Score 56; DB 10;
Pred. No. 6.3;
                                                                                                                                                                                                  ж
;
                                                                                                                                                                                                                            EKLFKEVKLEELENQKGSGLGKAQCAALWLQCASGG 37
                                                                                                                                                                                                  Score 61; DB Pred. No. 7.2;
                                                                                                                                                                                                                                                                                                                                                                                                                                               Mismatches
                                                                                                                                                                                                               Mismatches
                                                                                                                                  BY SIMILARITY
                                                                                                                                                                                                                                   LFKEVKLEELENQKGSGLGKAQCAAL--WL 31
                                                                                                                                                                                                                                                                                     Created)
                                                                                                                                                                                                                                                                         PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                               5,
                                                                                                                                                                                                               э;
                                                                                                                                                                                                                                                                          01-DEC-2001 (TrEMBLrel. 19, C 01-DEC-2001 (TrEMBLrel. 19, L 01-DEC-2001 (TrEMBLrel. 19, L UNKNOWN PROTEIN. AT2G21970.
                                                                                                                                                                                    88052 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                 29.0%;
illarity 43.3%;
Conservative
                                                                                                                                                                                                  31.6%;
36.1%;
                                                                                                                                                                                                               Conservative
                                                                                                                                                                                                                                                                        PRELIMINARY
  l Similarity
13; Conserv
  639
647
663
663
671
687
687
711
719
727
727
759
759
775
775
783
832
309
359
845
845
                                                                                                                                                                                                                                                                                                                                                                                                                                         Local Similarity
nes 13; Conser
                                                                                                                                                                                                                                                                                                                                       eurosids II; Bra
NCBI_TaxID=3702;
[1]
                                                                                                                                                                                                                                                                       Q94F34;
Q94F34;
O1-DEC-2001
O1-DEC-2001
                                                                                                                                                   CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
CARBOHYD
SEQUENCE
                                                                                                                                       CARBOHYD
CARBOHYD
                                                                                                                                                                                                                                                                                                                                                                                                                                  Query Match
                                                                                                                                  ULFID
                                                                                                                                                                                                  Query Match
                                                                                                                                                                                                         Local
                          REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
REPEAT
                                                                                          REPEAT
REPEAT
REPEAT
                                                                                                                    REPEAT
DOMAIN
        REPEAT
REPEAT
REPEAT
                                                                                                              REPEAT
                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                            4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                         86
                                                                                                                                                                                                                            ~
                                                                                                                                                                                                                                        801
                                                                                                                                                                                                              Matches
                                                                                                                                  DISI
                                                                                                                                                                                                                                                           4
                                                                                                                                                                                                         Best
                                                                                                                                                                                                                                                           RESULT
                                                                                                                                                                                                                                                                 94F34
                                                                                                                                                                                                                                                                              d
  a
                                                                                                                                                                                                                                                                         G
                                                                                                                                                                                                                            à
```

```
;
                                                                                                                                                                                                                                                                                                                                           SEQUENCE FROM N.A.
STRAIN=CV. COLUMBIA;
MEDLINE=20083487; PubMed=10617197;
Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
Salzberg S.L., Fraser C.M., Venter J.C.;
"Sequence and analysis of chromosome II of Arabidopsis thaliana.";
Nature 402:761-768(1999).
09SJ02 PRELIMINARY; PRT; 202 AA.
09SJ02;
01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
ATG21970 PROTEIN (STRESS ENHANCED PROTEIN 2).
ATG21970 OR SEP2.
Arabidopsis thaliana (Mouse-ear cress).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
NCBI_TaxID=3702;
[1]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
[1]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Heddad M., Adamska I.;
Heddad M., Adamska I.;
"Light stress-regulated two-helix proteins in Arabidopsis thaliana related to the chlorophyll a/b-binding gene family.";
Proc. Natl. Acad. Sci. U.S.A. 97:3741-3746(2000).
EMBL; AC007019; AAD20414.1; -.
EMBL; AF133717; AAF61626.1; -.
SEQUENCE 202 AA; 21989 MW; D300A5BIEE70F6F8 CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  "Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium."; Science 282:2012-2018(1998).
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      2
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   202;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                (TrEMBLrel. 15, Created)
(TrEMBLrel. 15, Last sequence update)
(TrEMBLrel. 19, Last annotation update)
L 59.2 KDA PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DB 10;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 56; DB Pred. No. 7.6;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   513
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               159
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              LFKEVKLEELENQKGSGLGKAQCAAL--WL 31
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            ||||:::|| |::||
|130 LFKKLDVEGLSEAIGAGLAAMGCAAMFAWL
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             PubMed=10725357;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
STRAIN-BRISTOL N2;
MEDLINE=99069613; PubMed=9851916;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   29.0%;
43.3%;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       SEQUENCE FROM N.A.
STRAIN=BRISTOL N2;
Wollam C., Zidanic M., Du
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        elegans
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Similarity
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               MEDLINE=20202706;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Q9N2W3 PREI
Q9N2W3;
01-OCT-2000 (TrE
01-OCT-2001 (TrE
01-DEC-2001 (TrE
HYPOTHETICAL 59.
Y94H6A.1.
```

```
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                            Query Match
                                                                                                                                                                                                                                                                                                                                                             Kinase
             416
                                                                                                                                                                                                                                                                                                                                                                                                      Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                    Н
                                                                                                                                                                                                                                                                                                                                                                                                                                                       416
                                             œ
                                             RESULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       RESULT
                                                       GN1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                09GN0
                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    δ
                                                                                                                                                                                                                                                                                     ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  comorpha;
                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          in
                                                                                                                                                                                                                                                                                    ö
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          es
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
SEQUENCE FROM N.A.
STRAIN=HFL97_4S, VA96_7S, GA96_2S, AND GA96_4S;
Duvernell D.D., Eanes W.F.;
"Contrasting molecular population genetics of four hexokinase Drosophila melanogaster and Drosophila simulans.";
Genetics 0:0-0(2000).
EMBL; AF257649; AAG22968.1; -.
EMBL; AF257644; AAG22954.1; -.
EMBL; AF257644; AAG22958.1; -.
EMBL; AF257645; AAG22960.1; -.
                                                                                                                                                                                                         PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
DNA-binding; Hypothetical protein; Nuclear protein; Receptor Transcription regulation; Zinc-finger.
SEQUENCE 513 AA; 59164 MW; C79EDF2C65E99A6E CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Drosophila simulans (Fruit fly).

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muse
Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7240;
                                                                                              FAM
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      453;
                                                                                                                                                                                                                                                                513;
                                                           "Direct Submission.";
Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
-!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
-!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS EMBL; AC024876; AAF60896.1; -.
                                                                                                                                                                                                                                                                                     Indels
cosmid Y94H6A.";
EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Indel
                                                                                                                                                                                                                                                                Length
                                                                                                                                                                                                                                                                                                                                                                                                            Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                   15;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                E530FBF0E717B821
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     5;
                                                                                                                                                                                                                                                               5;
                                                                                                                                                                                                                                                                                                                                                                                 ¥
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Score 54.5; D
Pred. No. 29;
5; Mismatches
                                                                                                                                                                                                                                                                DB
                                                                                                                                                                                                                                                                                   Mismatches
                                                                                                                                                                                                                                                                         20
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         MEKLFKE-VKLEELENQKGSGLGKAQCAALWLQCAS
                                                                                                                                                                                                                                                                                                                                                                               453
                                                                                                                                                                                                                                                               26;
                                                                                                                                                                                                                                                                         N<sub>O</sub>
                                                                                                                                                                                                                                                                                                                  311 FLKIQREEFENCPEDGQIPAPCSAEWFQELSG
                                                                                                                                                                                                                                                                                                        FKEVKLEELENQKGSGLGKAQCAALWLQCASG
                                                                                                              HSSP; P20393; 1A6Y.
InterPro; IPR000536; Hormone_rec_lig
InterPro; IPR001628; zf-C4.
InterPro; IPR000130; Zn_MTpeptdse.
                                                                                                                                                                                                                                                                                                                                                                                                   Created)
                                                                                                                                                                                                                                                                Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   FlyBase; FBgn0042855; Dsim\Hex-t2
InterPro; IPR001312; Hexokinase.
Pfam; PF00349; hexokinase: 1
                                                                                                                                                                                                                                                                                   4;
                                                                                                                                                                          PRINTS; PR00047; STROIDFINGER. SMART; SM00430; HOLI; 1. SMART; SM00399; ZnF_C4; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   PRINTS; PR00475; HEXOKINASE.
ProDom; PD001109; Hexokinase; 1
PROSITE; PS00378; HEXOKINASES;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               49856 MW;
                                                                                                                                                                                                                                                                                                                                                                                                 (TrEMBLrel. 16, 1 (TrEMBLrel. 16, 1 (TrEMBLrel. 19, 1
elegans
to the
                                                                                                                                                                                                                                                               29.0%;
llarity 40.6%;
Conservative
                                                                                                                                                            rec;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Similarity 44.4%; 6; Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           hexokinase;
                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                       Pfam; PF00104; hormone
Pfam; PF00105; zf-C4; 1
"The sequence of C. Submitted (MAR-2000)
                              SEQUENCE FROM N.A. STRAIN=BRISTOL N2;
                                                                                                                                                                                                                                                                        Local Similarity
les 13; Conser
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              453 AA;
                                                                                                                                                                                                                                                                                                                                                                                     09GN24;
01-MAR-2001 (Tr
01-MAR-2001 (Tr
01-DEC-2001 (Tr
HEXOKINASE-T2.
                                                   Waterston R.;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Query Match
Best Local S
Matches 16
                                                                                                                                                                                                                                                              Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                              09GN24
                                                                                                                                                                                                                                                                                                      S
                                                                                                                                                                                                                                                                                  Matches
                                                                                                                                                                                                                                                                                                                                                          RESULT
                                                                                                                                                                                                                                                                                                                                                                   Q9GN24
                                                 qq
                                                                                                                                                                                                                                                                                                                                                                                                ð
```

ŏ

```
Oggn13
PRELIMINARY;
PRT; 453 AA.

OgGN13;
O1-MAR-2001 (TrEMBLrel. 16, Last sequence update)
O1-MAR-2001 (TrEMBLrel. 19, Last annotation update)
O1-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HEXOKINASE-T2.
Drosophila simulans (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Brachycera; Muscomorpha;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                            in
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        :ics of four hexokinases in
simulans.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                          of four hexokinases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                AND VA96
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Insecta;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Q9GN05
Q9GN05;
Q9GN05;
Q1-MAR-2001 (TrEMBLrel. 16, Created)
O1-MAR-2001 (TrEMBLrel. 16, Last sequence update)
O1-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HEXOKINASE-T2.
Drosophila simulans (Fruit fly).
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Inserygota; Neoptera; Endopterygota; Diptera; Brachycera Ephydroidea; Drosophilae; Drosophila.
NCBI_TaxID=7240;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              CT96_6S,
                                                                                                                                                                                                                                                                                                                                                                                                    SEQUENCE FROM N.A.
STRAIN=HFL97_2S, VA96_8S, AND HFL97_1S;
Duvernell D.D., Eanes W.F.;
"Contrasting molecular population genetics of four Drosophila melanogaster and Drosophila simulans.";
Genetics 0:0-0(2000).
EMBL; AF257648; AAG22966.1; -.
EMBL; AF257647; AAG22966.1; -.
EMBL; AF257647; AAG22964.1; -.
HSSP; P19367; 1HKB.
FlyBase; FBgn0042855; Dsim\Hex-t2.
Interpro; IPR001312; Hexokinase.
Pfam; PF00349; hexokinase; 1.
PRINTS; PR00475; HEXOKINASE.
ProDom; PD001109; Hexokinase; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 EB93A6537F1B6504
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 14
                    451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    451
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           35
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CT96_5S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              ed. No. 29;
Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       lation genetics
Drosophila sim
54.5;
No. 29;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          MEKLFKE-VKLEELENQKGSGLGKAQCAALWLQCAS
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   SEQUENCE FROM N.A.
STRAIN=GA96_5S, DPF96_3S, CT96_2S, C
Duvernell D.D., Eanes W.F.;
"Contrasting molecular population ge
Drosophila melanogaster and Drosophi
Genetics 0:0-0(2000).
EMBL; AF257646; AAG22962.1; -.
EMBL; AF257637; AAG22944.1; -.
EMBL; AF257639; AAG22946.1; -.
EMBL; AF257640; AAG22946.1; -.
EMBL; AF257640; AAG22950.1; -.
EMBL; AF257641; AAG22950.1; -.
EMBL; AF257641; AAG22950.1; -.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               CT96_2S,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               2;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 49843 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28.28;
44.48;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Local Similarity 44.4
                                                                                                                                                                                                                                                                                                                                                                                            SEQUENCE FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 453 AA;
```

1;

2002

14:40:04

24

Oct

Thu

•

1;

```
Similarity
16; Conser
                                                                                        Similarity
                                                                                                                                                                                                           O9NY22;
01-OCT-2000
01-OCT-2001
01-DEC-2001
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 ANK repeat;
                                                                                                 16;
                                                                                                                                                                                                                                                                                                                                                                                                                                                              InterPro;
Pfam; PF0
Pfam; PF0
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             Query Match
Best Local S
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE
                                                                             Query Match
                                                                                       Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Q9NQ74
                                                                                                                                                                                                 09NY22
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Best Loc
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             319
                                                                                                                        m
                                                                                                                                            319
                                                                                                 Matches
                                                                                                                                                                            12
                                                                                                                                                                                                                                                                 VR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   VR1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             RESULT
Q9NQ74
                                                                                                                                                                            RESULT
09NY 22
                                                                                        Best
                                                                                                                                                                                                                                                                DE DE OC OC OC OC OC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ΩI
   DR
DR
DR
SQ
                                                                                                                                                                                                                      DT DT DE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      담
                                                                                                                                            qq
                                                                                                                        ğ
                                                                                                                        ;;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               1;
                                                                                                                                                                                                                                                                                                                                                               VR1.";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Gaps
                                                                                                                        Gaps
                                                                                                                                                                                                                                                                                                  Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Vertebrata; Euteleostomi; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                 tor,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        tor,
                                                                                                                        1;
                                                                                                                                                                                                                                                                                                                               SEQUENCE FROM N.A.

Cortright D.N., Peck A., Zou T.;

Cloning and characterization of the human capsaicin receptor,

Expense;

Rembl, AF196176; AAG43467.1; -.

InterPro; IPR002110; ANK.

R SMART; SM00248; ANK,

R SMART; SM00248; ANK_REPEAT; 1.

PROSITE; PS50297; ANK_REP_REGION; 1.

PROSITE; PS50297; ANK_REP_REGION; 1.

ANK repeat; Receptor; Repeat.

ANK repeat; Receptor; Repeat.

Though 511;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Cortright D.N., Peck A., Zou T.;
Cortright D.N., Peck A., Zou T.;
"Cloning and characterization of the human capsaicin recept
Submitted (OCT-1999) to the EMBL/GenBank/DDBJ databases.
EMBL; AF196175; AAG43466.1; -.
InterPro; IPR002110; ANK.
InterPro; IPR000636; Cation_chan_non_lig.
InterPro; IPR002111; Cat_channel_TrpL.
                                                                                                  45
                                                                                                                                                                                                                                                                                     Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                        Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               Indels
                                                                                                  Length
                                                                            FE87E3537F1B651F CRC64;
                                                                                                                                                                                                                                                       Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            update)
                                                                                                                      14;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               11;
                                                                                                  5;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     32
                                                                                                                                                                  451
                                                                                                   DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              839 AA
                                                                                                 Score 54.5; D
Pred. No. 29;
5; Mismatches
                                                                                                                                             m
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    33;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -- SGLGKAQCAALWLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Ve
Mammalia; Eutheria; Primates; Catarrhini;
NCBI_TaxID=9606;
                                                                                                                                                         |:|| | || || || || || || || || || 416 MKKLLKPGVKFELIVSEDGSGRGAALVAATAVQAKS
                                                                                                                                            MEKLFKE-VKLEELENQKGSGLGKAQCAALWLQCAS
                                                                                                                                                                                                                        511
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 54.5;
Pred. No. 3
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Created)
Last sequ
                                                                                                                                                                                                                                            Created)
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               PRT;
                                                                                                                                                                                                                       PRT;
FlyBase; FBgn0042855; Dsim\Hex-t2.
InterPro; IPR001312; Hexokinase.
Pfam; PF00349; hexokinase; 1.
PRINTS; PR00475; HEXOKINASE.
ProDom; PD001109; Hexokinase; 1.
PROSITE; PS00378; HEXOKINASES; 1.
   Dsim\Hex-t2
                                                                                                                                                                                                                                                                 Last
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               3;
                                                                                                                      5
                                                                            49869 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                        28.2%;
illarity 43.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 16, (TrEMBLrel. 16, 16, 17)
                                                                                                28.2%;
ilarity 44.4%;
Conservative
                                                                                                                                                                                                                                          01-MAR-2001 (TrEMBLrel. 16, 01-MAR-2001 (TrEMBLrel. 16, 01-DEC-2001 (TrEMBLrel. 19, CAPSAICIN RECEPTOR VARIANT.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                       PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    3 KLFKEVKLEELENQKG-
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     09H304;
01-MAR-2001 (TrEMBLE
01-MAR-2001 (TrEMBLE
01-DEC-2001 (TrEMBLE
CAPSAICIN RECEPTOR.
                                                                                                          l Similarity
16; Conserv
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Local Similarity
nes 16; Conser
                                                                            453 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SEQUENCE FROM N.A
                                                                                                                                                                                                                                                                                                                    NCBI_TaxID=9606;
                                                                                                                                                                                                                    Q9H3O3
Q9H3O3;
01-MAR-2001
                                                                  Kinase.
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Query Match
                                                                                                  Query Match
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Q9H304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        319
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Matches
                                                                                                                                             Н
                                                                                                                      Matches
                                                                                                                                                                                                 RESULT 10
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        RESULT
Q9H304
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SULT
                                                                                                                                                                                                            09H30
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   DR DR DR DR SO SO
                                                                                                                                                                                                                                                    d
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  RP
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           RA
                                                                                                                                                                  g
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     δ
                                                                                                                                             ð
```

```
McIntyre P., McLatchie L., Chambers A., Phillips E., Clarke M., Savidge J., Peacock M., Shah K., Winter J., Weerasekera N., Webb M., Rang H., Bevan S., James I.; "Pharmacological differences between human and rat vanilloid receptor 1 (VR1).";
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    Gaps
                                                                                                                                                             Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Euteleostomi;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   7;
                                                                                                                                                           7;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      839;
                                                                                                                              839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Craniata; Vertebrata; Eutel
Catarrhini; Hominidae; Homo
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                              Length
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Length
                                                                                                                                                             Indel
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        773BF59D43189968 CRC64;
                                                                                               CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Last sequence update)
Last annotation update)
                                                                                                                                                                                                                                                                                                                                                             update)
                                                                                                                                                                                                                                                                                                                                                 update)
                                                                                                                                                             11;
                                                                                                                            DB 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   11
                                                                                               EA9BF7D6023EF9F4
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  355
                                                                                                                                                                                                                        355
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    32
                                                                                                                                                                                            32
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        DB
                                                                                                                                                                                                                                                                                                                                  Created)
Last sequence up
Last annotation
                                                                                                                                                                                                                                                                                                     AA
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 54.5; D. Pred. No. 56; 3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                Æ
                                                                                                                Score 54.5; Di
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   -SGLGKAQCAALWLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    -SGLGKAQCAALWLQ
                                                                                                                                                                                                            Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                l (VR1).";
Br. J. Pharmacol. 132:1084-1094(2001).
EMBL; AJ272063; CAB89866.2; -.
InterPro; IPR002110; ANK.
InterPro; IPR000636; Cation_chan_non_lig
InterPro; IPR002111; Cat_channel_TrpL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             839
                                                                                                                                                                                                                                                                                                     839
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         SMART; SM00248; ANK; 4.
PROSITE; PS50088; ANK_REPEAT; 1.
PROSITE; PS50297; ANK_REP_REGION; 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            Created)
Pfam; PF00520; ion_tid....
SMART; SM00248; ANK; 4.
PROSITE; PS50088; ANK_REPEAT; 1.
PROSITE; PS50297; ANK_REP_REGION; 1
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRT;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          SEQUENCE FROM N.A.
TISSUE=DORSAL ROOT GANGLION;
MEDLINE=21124568; PubMed=11226139;
                                                                                                                                                           3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        Homo sapiens (Human).
Eukaryota; Metazoa; Chordata;
Mammalia; Eutheria; Primates;
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          .
Έ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Receptor; Repeat 839 AA; 94968 MW
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     28.2%;
ilarity 43.2%;
Conservative
                                                                                                                          28.2%;
illarity 43.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          PF00023; ank; 3.
PF00520; ion_trans; 1.
; SM00248; ANK; 4.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           15,
15,
19,
                                                                                                                                                                                                                                                                                                                                  (TrEMBLrel. 15, (TrEMBLrel. 18, (TrEMBLrel. 19,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Q9NQ74;
01-OCT-2000 (TrEMBLrel. 1
01-OCT-2000 (TrEMBLrel. 1
01-DEC-2001 (TrEMBLrel. 1
VANILLOID RECEPTOR 1.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  KLFKEVKLEELENQKG--
                                                                                                                                                                                                                                                                                                     PRELIMINARY
                                                                                                                                                                                            KLFKEVKLEELENQKG-
                                                                                                                                                                                                                                                                                                                                                                                                               Homo sapiens (Human)
                                                                                                                                                                                                                                                                                                                                                                                VANILLOID RECEPTOR 1
```

.; H .rspt

37

us-09-894-030-3\_copy\_l

```
NON_TER
SEQUENCE
                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Local
                                                  319
                          m
                                                                                          RESULT 15
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Matches
Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Search co
Job time
                                                                                                                                                                                                                                                                                                                                                                                                                                                                       Best
                                                                                                        O9IEZ2
                                                                                                                                              qq
                                                                                                                                                                                                                                                                                                                                                                            ij
                                              SEQUENCE FROM N.A.

MEDLINE=20504379; PubMed=11050376;
Hayes P., Meadows H.J., Gunthorpe M., Harries M.H., Duckworth M.D.,
Cairns W., Harrison D.C., Clarke C., Ellington K., Prinjha R.K.,
Barton A.J., Medhurst A.D., Smith G.D., Topp S., Murdock P.,
Sanger G.J., Terrett J., Jenkins O., Benham C.D., Randall A.D.,
Gloger I.S., Davis J.B.;
"Cloning and functional expression of a human orthologue of rat
vanilloid receptor-1.";
Pain 88:205-215(2000).
REMBL; AJ277028; CAB95729.1; -.
RINTERPRO; IPR002110; ANK.
RINTERPRO; IPR00636; Cation_chan_non_lig.
RINTERPRO; IPR00520; ion_trans; 1.
                                                                                                                                                                                                                                                                                                                                                                           Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              SEQUENCE FROM N.A.

TISSUE=TESTIS;

MEDLINE=21154917; PubMed=11230166;

Miemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,

Ansorge W., Boecher M., Bloecker H., Bauersachs S., Blum H.,

Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,

Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,

Wambutt R., Korn B., Klein M., Poustka A.;

"Towards a Catalog of Human Genes and Proteins: Sequencing and
"Tamalysis of 500 Novel Complete Protein Coding Human cDNAs.";

Genome Res. 11:422-435(2001).

REMBL; AL136801; CAB66735.1; -.

InterPro; IPR002110; ANK.

InterPro; IPR002111; Cat_channel_TrpL.

InterPro; IPR002111; Cat_channel_TrpL.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Homo sapiens (Human).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    s:
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          D.,
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   and
                                                                                                                                                                                                                                                                                                                                               Length 839;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 839;
                                                                                                                                                                                                                                                                                                                                                                           Indels
                       Kelsell R.E.;
Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              Pfam; rrvv--
SMART; SM00248; ANK; 4.
PROSITE; PS50088; ANK_REPEAT; 1.
PROSITE; PS50297; ANK_REP_REGION; 1.
PROSITE; PS50297; ANK_REP_REGION; 1.
                                                                                                                                                                                                                                                                                                                    7142FFAE43189ECC CRC64;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       09H0G9;
01-MAR-2001 (TrEMBLrel. 16, Created)
01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
HYPOTHETICAL 94.9 KDA PROTEIN.
                                                                                                                                                                                                                                                                                                                                                                         11;
                                                                                                                                                                                                                                                                                                                                                 4;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          4;
                                                                                                                                                                                                                                                                                                                                                 DB
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           DB
                                                                                                                                                                                                                                                                                                                                               Score 54.5; Di
Pred. No. 56;
3; Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                               --SGLGKAQCAALWLQ
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        54.5; I
No. 56;
                                                                                                                                                                                                                                          InterFIG; IFROCEALT, CONTROL Pfam; PF00520; ion_trans; 1.
SMART; SM00248; ANK; 3.
PROSITE; PS50088; ANK_REPEAT; 1.
PROSITE; PS50297; ANK_REP_REGION; 1
ANK repeat; Receptor; Repeat.
SFOUENCE 839 AA; 94924 MW; 7142]
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          Score
Pred.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRT;
                                                                                                                                                                                                                                                                                                                                                                        3;
                                                                                                                                                                                                                                                                                                                                               28.2%;
llarity 43.2%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         28.28;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        FUU520; ion_trans; 1
SMO0248; ANK; 4.
                                                                                                                                                                                                                                                                                                                                                                                                3 KLFKEVKLEELENQKG----
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            PRELIMINARY;
                                                                                                                                                                                                                                                                                                                                                           Local Similarity
tes 16; Conser
          FROM N.A.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Eukaryota; Metazo
Mammalia; Eutheri
NCBI_TaxID=9606;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 SMART; SMOOT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          DKFZP434K0220.
        SEQUENCE
                                                                                                                                                                                                                                                                                                                                               Query Match
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Ratch
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                         Query Maich
Best Local
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              09нод
                                                                                                                                                                                                                                                                                                                                                                        Matches
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    14
                                                                                                                                                                                                                                                                                                                                                            Best
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    SULT
                                                                                                                                                                                                                                                                                                                                                                                                                                                                               бэнбб
                                                                                                    RAA
RAA
RAA
RET
RET
DR BDR
DR DR DR
DR KW
KW
KW
SQ
                                                                                                                                                                                                                                                                                                                                                                                                                             d
                                                                                                                                                                                                                                                                                                                                                                                                    δ
```

Similarity

```
1;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     ö
 Gaps
                                                                                                                                                                                                                                                                                                                                                                                                                                                                      Gaps
                                                                                                                                                                                                                                                     Valas S., Benoit C., Baudry C., Perrin G., Mamoun R.Z.; "Variability and immunogenicity of Caprine arthritis-encephalitis virol. 74:6178-6185(2000).

EMBL; AJ400719; CAB95712.1; -. Signal.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    ö
~
                                                                                                                                                                                                                                                                                                                                                                                                                                           Length 634;
 Indels
                                                                                                                                                                                                                                                                                                                                                                                                                                                                     Indels
                                                                                                                                                                                                                                                                                                                                                                                                                 75B196E23560F23E CRC64;
                                                                                                               Q9IEZ2;
Q9IEZ2;
Q9IEZ2;
01-OCT-2000 (TrEMBLrel. 15, Created)
01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
SURFACE GLYCOPROTEIN PRECURSOR (FRAGMENT).
ENV.
                                                                                                                                                                                                       Caprine arthritis encephalitis virus (CAEV).
Viruses; Retroid viruses; Retroviridae; Lentivirus
NCBI_TaxID=11660;
                                                                                                                                                                                                                                                                                                                                                                            POTENTIAL. SURFACE GLYCOPROTEIN.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                    8;
11;
                                                                                                                                                                                                                                                                                                                                                                                                                                        Score 54; DB 15;
Pred. No. 48;
; Mismatches 8;
                                                   355
                          32
                        Mismatches
                                                                                                                                                                                                                                                                                                                                                                                                                                     Score 54;
Pred. No.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           09:29:16
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             2 EKLFKEVKLEELENQKGSGLGKAQ 25
                                                                                                                                                                                                                                                                                                                                                                                                                                                                  4;
3;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            2002,
                                                                                                                                                                                                                                                                                                                                                                           86 PP >634 S 634 S 172847 MW;
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           || :||| :|| | :|| | EKNWKEVIIEEEEEKKGEGCTRRQ
                                                                                                                                                                                                                                                                                                                                                                                                                                          28.0%;
ilarity 50.0%;
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            24,
Conservative
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            October
                                                                                                                                                                                                                                                [1]
SEQUENCE FROM N.A.
STRAIN=021;
MEDLINE=20305014; I
                                                                                                                                                                                                                                                                                                                                                                                                                 634 AA;
                                                                                                                                                                                                                                                                                                                                                                                                                                                       Similarity 12; Conser
                                                                                                                                                                                                                                                                                                                                                                             1
87
634
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            completed: (
16;
```